High-resolution scalar quantization with Rényi entropy constraint

Wolfgang Kreitmeier and Tamás Linder

July 5, 2011

Abstract

We consider optimal scalar quantization with rth power distortion and constrained Rényi entropy of order α. For sources with absolutely continuous distributions the high rate asymptotics of the quantizer distortion has long been known for $\alpha = 0$ (fixed-rate quantization) and $\alpha = 1$ (entropy-constrained quantization). These results have recently been extended to quantization with Rényi entropy constraint of order $\alpha \geq r+1$. Here we consider the more challenging case $\alpha \in [-\infty, 0) \cup (0, 1)$ and for a large class of absolutely continuous source distributions we determine the sharp asymptotics of the optimal quantization distortion. The achievability proof is based on finding (asymptotically) optimal quantizers via the companding approach, and is thus constructive.

Index Terms: companding, high-resolution asymptotics, optimal quantization, Rényi entropy.

1 Introduction

With the exception of a few very special source distributions the exact analysis of the performance of optimal quantizers is a notoriously hard problem. The asymptotic theory of quantization facilitates such analyses by assuming that the quantizer operates at asymptotically high rates. The seminal work by Zador [31] determined the asymptotic behavior of the minimum quantizer distortion under a constraint on either the log-cardinality of the quantizer codebook (fixed-rate quantization) or the Shannon entropy of the quantizer output (entropy-constrained quantization). (See the article by Gray and Neuhoff [12] for a historical overview and related results.) Zador's results were later clarified and generalized by Bucklew and Wise [5] and Graf and Luschgy [9] for the fixed-rate case, and by Gray *et al*. [11] for the entropy-constrained case.

Recently, approaches that incorporate both the fixed and entropy-constrained cases have been suggested. In [10] a Lagrangian formulation is developed which puts a simultaneous constraints on entropy and codebook size, including fixed-rate and entropy-constrained quantization as special cases. Another

W. Kreitmeier is with the Department of Informatics and Mathematics, University of Passau, Innstraße 33, 94032 Passau, Germany (email: `wolfgang.kreitmeier@uni-passau.de`). T. Linder is with the Department of Mathematics and Statistics, Queen's University, Kingston, Ontario, Canada K7L 3N6 (email: `linder@mast.queensu.ca`).

This research was supported in part by the German Research Foundation (DFG) and the Natural Sciences and Engineering Research Council (NSERC) of Canada.

approach that has been suggested in [10] and further developed in [18, 19] uses the Rényi entropy of order α of the quantizer output as (generalized) rate. One obtains fixed-rate quantization for $\alpha = 0$, while $\alpha = 1$ yields the usual (Shannon) entropy-constrained quantization approach.

The choice of Rényi entropy as the quantizer's rate can be motivated from a purely mathematical viewpoint. In the axiomatic approach to defining entropy, Rényi's entropy is a canonical extension of Shannon-entropy, satisfying fewer of the entropy axioms [26, 1]. From a more practical point of view, the use of Rényi entropy as quantizer rate is supported by Campbell's work [6], who considered variable-length lossless codes with exponentially weighted average codeword length and showed that Rényi's entropy plays an analogous role to Shannon entropy in this more general setting. Further results on lossless coding for Rényi entropy were obtained in [24]. Jelinek [16] showed that Rényi's entropy (of an appropriate order $\alpha \in (0, 1)$) of a variable-length lossless code determines the encoding rate for a given reliability (exponential decrease of probability) of buffer overflow when the codewords are transmitted over a noiseless channel at a fixed per symbol rate. At least in such situations, measuring the quantizer's rate by Rényi's entropy is operationally justified. An overview of related results can be found in [2]. The diverse uses of Rényi's entropy (and differential entropy) in emerging fields such as quantum information theory (e.g. [15]), statistical learning (e.g. [17]), bioinformatics (e.g. [21]), etc., may also provide future motivation for this rate concept.

The only available general result on quantization with Rényi entropy constraint appears to be [18] where the sharp asymptotic behavior of the rth power distortion of optimal d-dimensional vector quantizers has been derived for $\alpha \in [1 + r/d, \infty]$. The proof shows that for these α values the optimal quantization error is asymptotically determined by the distortion of a ball with appropriate radius around the most likely values of the source distribution. Thus it suffices to evaluate the rth moment of this ball (see [18, Theorem 4.3]), which remarkably simplifies the derivation and makes the case $\alpha \geq 1 + r/d$ quite unique. In the classical ($\alpha = 0$ and $\alpha = 1$) settings, the contributions of the codecells of an optimal quantizer to the overall distortion are asymptotically of the same order. Bounds on the optimal performance in [18] suggest a similar situation for $\alpha < 1 + r/d$, making the problem more challenging than the case $\alpha \geq 1 + r/d$.

In this paper, at the price of restricting the treatment to the scalar ($d = 1$) case, we are able to determine the asymptotics of the optimal quantization error under a Rényi entropy constraint of order $\alpha \in [-\infty, 0) \cup (0, 1)$ for a fairly large class of source densities. The achievability part of the proof (providing a sharp upper bound on the asymptotic performance) is constructive via companding quantization. In particular, we determine the optimal point density function for each $\alpha \in [-\infty, 1 + r)$ and provide rigorous performance guarantees for the associated companding quantizers (for $\alpha = 0$ and $\alpha = 1$, these results have of course been known). Matching lower bounds are provided for $\alpha \in [-\infty, 0) \cup (0, 1)$, which leaves only the case $\alpha \in (1, 1 + r)$ open. We note that in proving the matching lower bounds, one cannot simply apply the techniques established for $\alpha = 0$ or $\alpha = 1$. In our case the distortion and Rényi entropy of a quantizer must be simultaneously controlled, a difficulty not encountered in fixed-rate quantization. Similarly, the Lagrangian formulation that facilitated the corrected proof of Zador's entropy-constrained

quantization result in [11] cannot be used since it relies on the special functional form of the Shannon entropy. On the other hand, using the monotonicity in α of the optimal quantization error, one can show that our results imply the well-known asymptotics for $\alpha \in \{0, 1\}$, at least for the special class of scalar distributions we consider.

The paper is organized as follows. In Section 2 we introduce the quantization problem under a Rényi entropy constraint and review some definitions and notation. In Section 3, after summarizing some related work, we state our main result. The next three sections are devoted to developing the machinery needed in the proof. Section 4 presents results on the asymptotic distortion and Rényi entropy of companding quantizers, which, with the proper choice of the compressor function in a Bennett-like integral, will turn out to be (asymptotically) optimal. In Section 5 technical results needed mostly for establishing lower bounds are developed. Section 6 presents upper and lower bounds on the optimal quantization error for mixture distributions. Section 7 contains the proof of the main results. Section 8 contains concluding remarks and a discussion on extending the results to vector quantization. All the longer, technical proofs of the auxiliary results are relegated to the appendices.

2 Preliminaries and notation

We begin with the definition of Rényi entropy of order α.

Definition 2.1. *Let $\mathbb{N} := \{1, 2, \ldots\}$. Let $\alpha \in [-\infty, \infty]$ and $p = (p_1, p_2, \ldots) \in [0, 1]^{\mathbb{N}}$ be a probability vector, i.e., $\sum_{i=1}^{\infty} p_i = 1$. The Rényi entropy of order α, $\hat{H}^\alpha(p) \in [0, \infty]$, is defined as (see [26], [1, Definition 5.2.35] and [14, p. 1])*

$$\hat{H}^\alpha(p) = \begin{cases} \frac{1}{1-\alpha} \log \left(\sum\limits_{i:p_i>0} p_i^\alpha \right), & \alpha \in (-\infty, \infty) \setminus \{1\} \\ -\sum\limits_{i=1}^{\infty} p_i \log p_i, & \alpha = 1 \\ -\log\left(\max\{p_i : i \in \mathbb{N}\}\right), & \alpha = \infty \\ -\log\left(\inf\{p_i : i \in \mathbb{N}, p_i > 0\}\right), & \alpha = -\infty. \end{cases}$$

We use the conventions $0 \cdot \log 0 := 0$ and $0^0 := 0$. All logarithms are to the base e.

Remark 2.2. (a) With these conventions we obtain

$$\hat{H}^0(p) = \log\left(\mathrm{card}\{i \in \mathbb{N} : p_i > 0\}\right),$$

where card denotes cardinality. Using l'Hospital's rule it is easy to see, that the case $\alpha = 1$ follows from the case $\alpha \neq 1$ by taking the limit $\alpha \to 1$. (see, e.g., [1, Remark 5.2.34]). Moreover, one has

$$\lim_{\alpha\to\infty} \hat{H}^\alpha(p) = \hat{H}^\infty(p) \quad \text{and} \quad \lim_{\alpha\to-\infty} \hat{H}^\alpha(p) = \hat{H}^{-\infty}(p). \tag{1}$$

(b) We note that the usual definition of Rényi entropy is restricted to nonnegative values of the order α. However, it will turn out that the case $\alpha < 0$ can be handled without too much additional technical difficulties, and we believe that this generalization may turn out to have useful implications.

Now let $d \in \mathbb{N}$ and X be an $\mathbb{R}^d$-valued random variable with distribution μ. Let $\mathbb{I} \subset \mathbb{N}$ and $\mathcal{S} = \{S_i : i \in \mathbb{I}\}$ be a countable and Borel measurable partition of $\mathbb{R}^d$. Moreover let $\mathcal{C} = \{c_i : i \in \mathbb{I}\}$ be a countable set of distinct points in $\mathbb{R}^d$. Then $(\mathcal{S}, \mathcal{C})$ defines a *quantizer* $q : \mathbb{R}^d \to \mathcal{C}$ such that

$$q(x) = c_i \qquad \text{if and only if} \qquad x \in S_i.$$

We call $\mathcal{C}$ the *codebook* and the c_i the codepoints. Each $S_i \in \mathcal{S}$ is called *codecell*. Clearly, $\mathcal{C} = q(\mathbb{R}^d)$ (the range of q). Moreover,

$$\mathcal{S} = \{q^{-1}(z) : z \in q(\mathbb{R}^d)\}$$

where $q^{-1}(z) = \{x \in \mathbb{R}^d : q(x) = z\}$. Let $\mathcal{Q}_d$ denote the set of all quantizers on $\mathbb{R}^d$, i.e., the set of all Borel-measurable mappings $q : \mathbb{R}^d \to \mathbb{R}^d$ with a countable number of codepoints $q(\mathbb{R}^d)$. The discrete random variable $q(X)$ is a quantized version of the random variable X whose distribution is denoted by $\mu \circ q^{-1}$. In measure-theoretical terms the image measure $\mu \circ q^{-1}$ has a countable support and defines an approximation of μ, the so-called quantization of μ by q. With any enumeration $\{i_1, i_2, \ldots\}$ of $\mathbb{I}$ we define

$$H_\mu^\alpha(q) = \hat{H}^\alpha(\mu(S_{i_1}), \mu(S_{i_2}), \ldots) \tag{2}$$

as the Rényi entropy of order α of q with respect to μ. We intend to quantify the error in approximating the original distribution μ with its quantized version $\mu \circ q^{-1}$. To this end let $\|\cdot\|$ be any norm on $\mathbb{R}^d$ and $\rho : [0, \infty) \to [0, \infty)$ a strictly increasing function. For $q \in \mathcal{Q}_d$ we measure the approximation error between X and $q(X)$, resp. μ and $\mu \circ q^{-1}$, also called the quantizer distortion, as

$$D_\mu(q) = E\rho(\|X - q(X)\|) = \int \rho(\|x - q(x)\|)\, d\mu(x).$$

For any $R \geq 0$ we define

$$D_\mu^\alpha(R) = \inf\{D_\mu(q) : q \in \mathcal{Q}_d, H_\mu^\alpha(q) \leq R\}, \tag{3}$$

the optimal quantization distortion of μ under Rényi α-entropy bound R. We note that $D_\mu^\alpha(R)$ is a nonincreasing function of α (see Lemma 2.3).

We call a quantizer q optimal for μ under the entropy constraint R if $D_\mu(q) = D_\mu^\alpha(R)$ and $H_\mu^\alpha(q) \leq R$. In the rest of this paper we focus on the one-dimensional case (scalar quantizers, $d = 1$) and the so-called rth power distortion measure $\rho(x) = x^r$, where $r \geq 1$. Thus the distortion of quantizer $q \in \mathcal{Q}_1$ is given by

$$D_\mu(q) = E|X - q(X)|^r = \int |x - q(x)|^r\, d\mu(x).$$

For simplicity we write $\mathcal{Q}_1 = \mathcal{Q}$. Also, let $\mathcal{Q}^c \subset \mathcal{Q}$ denote the set of all scalar quantizers with finitely many codecells, each of which is an interval, and such that every codepoint lies in the closure of the corresponding codecell. The following lemma (proved in Appendix A) presents two key properties of optimal quantization under Rényi entropy constraint.

Lemma 2.3. *For all $R \geq 0$ and $\alpha, \beta \in [-\infty, \infty]$ with $\beta \leq \alpha$, we have*

$$D_\mu^\beta(R) \geq D_\mu^\alpha(R). \tag{4}$$

Assume that $E|X|^r < \infty$ and μ is nonatomic. Then for all $R \geq 0$ and $\alpha \in [-\infty, 0]$, we have

$$D_\mu^\alpha(R) = \inf\{D_\mu(q) : q \in \mathcal{Q}^c, H_\mu^\alpha(q) \leq R\} \tag{5}$$

while for all $\alpha \in (0, \infty]$,

$$D_\mu^\alpha(R) = \inf\{D_\mu(q) : q \in \mathcal{Q}^c, H_\mu^\alpha(q) = R\}. \tag{6}$$

The second statement of the lemma says that under the given conditions the optimum quantizer performance can be approached arbitrarily closely by quantizers in $\mathcal{Q}^c$. For this reason, in the rest of the paper all quantizers will be assumed to belong to $\mathcal{Q}^c$; in particular, we only consider quantizers with finitely many interval cells. According to (6), when $\alpha \in (0, \infty]$ it suffices to consider only those quantizers in $\mathcal{Q}^c$ whose entropy attains R.

From [18, Thm. 5.2] it is known that for $\alpha \in [0, 1]$ the product $e^{rR} D_\mu^\alpha(R)$ remains bounded and is bounded away from zero as $R \to \infty$. This motivates the following notion of quantizer optimality that will play an important role in our work.

Definition 2.4. *Let $(q_n)_{n\in\mathbb{N}} \subset \mathcal{Q}$ be a sequence of quantizers such that $H_\mu^\alpha(q_n) \to \infty$ as $n \to \infty$. If $e^{rR} D_\mu^\alpha(R) \to c$ as $R \to \infty$ for some $c \in (0, \infty)$ and*

$$\lim_{n\to\infty} e^{rH_\mu^\alpha(q_n)} D_\mu(q_n) = c, \tag{7}$$

then we call $(q_n)_{n\in\mathbb{N}}$ an asymptotically optimal sequence of quantizers for μ.

We denote by λ the one-dimensional Lebesgue measure. For a measurable real function f on $\mathbb{R}$ and measurable nonempty set $A \subset \mathbb{R}$, $\operatorname{ess\,inf}_A f = \sup\{b : \lambda(\{x \in A : f(x) < b\}) = 0\}$ denotes that the essential infimum of f on A. Similarly, $\operatorname{ess\,sup}_A f = \inf\{b : \lambda(\{x \in A : f(x) > b\}) = 0\}$ is the essential supremum of f on A. We let $\operatorname{supp}(\mu)$ denote the support of μ defined by

$$\operatorname{supp}(\mu) = \{x : \mu((x - \epsilon, x + \epsilon)) > 0 \text{ for all } \epsilon > 0\}.$$

Note that $\operatorname{supp}(\mu)$ is the smallest closed set whose complement has μ measure zero. We will often deal with the situation where $\operatorname{supp}(\mu)$ is contained in a bounded interval I. In such cases, we usually leave a quantizer $q \in \mathcal{Q}$ undefined outside I, as we may since $\mu(\mathbb{R} \setminus I) = 0$.

Let $\mathbb{Z}$ denote the set of all integers and assume $\Delta > 0$. The infinite-level uniform quantizer $\hat{q}_\Delta$ on $\mathbb{R}$ has codecells $\{(i\Delta, (i+1)\Delta] : i \in \mathbb{Z}\}$ and corresponding codepoints that are the midpoints of the associated cells, so that $\hat{q}_\Delta(x) = (i+1/2)\Delta$ if and only if $x \in (i\Delta, (i+1)\Delta]$.

3 Main results

First we summarize the known results regarding the sharp high-rate asymptotics of the distortion of optimal scalar quantizers. In order to unify the treatment, we reformulate the classical (resolution and entropy) rate constraints in terms of the Rényi entropy with appropriate order. For $r > 0$ we let

$$C(r) = \frac{1}{(1+r)2^r}.$$

Theorem 3.1 ([31, 5, 9, 11, 18]). *Let $r \geq 1$ and $\mu = \mu_a + \mu_s$ be the Lebesgue decomposition of distribution μ of the scalar random variable X with respect to the one-dimensional Lebesgue measure λ, where μ_a denotes the absolutely continuous part and μ_s the singular part of μ. Assume that $\mu_a(\mathbb{R}) > 0$ and let $f = \frac{d\mu_a}{d\lambda}$ be the density of μ_a.*

(i) If $\alpha = 0$ and $E|X|^{r+\delta} < \infty$ for some $\delta > 0$, then

$$\lim_{R\to\infty} e^{rR} D^0_\mu(R) = C(r) \left(\int f^{1/(1+r)}\, d\lambda \right)^{1+r}. \tag{8}$$

(ii) If $\alpha = 1$, $\mu_s(\mathbb{R}) = 0$, $\int f \log f\, d\lambda$ exists and is finite, and $H^1_\mu(\hat{q}_\Delta) < \infty$ for some $\Delta > 0$, then

$$\lim_{R\to\infty} e^{rR} D^1_\mu(R) = C(r) e^{-r\int f \log f\, d\lambda}. \tag{9}$$

(iii) If $\alpha \in [1+r, \infty]$, $\mu_s(\mathbb{R}) = 0$, $E|X|^{r+\delta} < \infty$ for some $\delta > 0$, and $\operatorname{ess\,sup}_{\mathbb{R}} f < \infty$, then

$$\lim_{R\to\infty} e^{(1+r)\beta(\alpha)R} D^\alpha_\mu(R) = C(r) \left(\operatorname{ess\,sup}_{\mathbb{R}} f\right)^{-r},$$

where $\beta(\alpha) = (\alpha-1)/\alpha$ if $\alpha \in [1+r, \infty)$ and $\beta(\alpha) = 1$ if $\alpha = \infty$.

Note that f is a probability density function if and only if $\mu_s(\mathbb{R}) = 0$. Part (i) of the theorem is originally due to Zador [31] who considered the multidimensional case; corrected and generalized proofs were given by Bucklew and Wise [5] and Graf and Luschgy [9]. Part (ii) is also due to Zador [31] with corrections and generalizations by Gray *et al*. [11]. Part (iii) is due to Kreitmeier [18] who also gave upper and lower bounds for the case $\alpha \in (1, 1+r)$.

Definition 3.2. *A one-dimensional probability density function f is called weakly unimodal if f is continuous on its support and there exists an $l_0 > 0$ such that $\{x : f(x) \geq l\}$ is a compact interval for every $l \in (0, l_0)$.*

Remark 3.3. Note if f is weakly unimodal density, then it is bounded and its support is a (possibly unbounded) interval. Clearly, all continuous unimodal densities are weakly unimodal. Thus the class of weakly unimodal densities includes most parametric source density classes commonly used in modeling information sources such as exponential, Laplacian, Gaussian, and generalized Gaussian densities.

For $\alpha \in (-\infty, r+1) \setminus \{1\}$ we define

$$a_1 = \frac{1-\alpha+\alpha r}{1-\alpha+r}, \qquad a_2 = \frac{1-\alpha+r}{1-\alpha}. \tag{10}$$

The following is the main result of the paper.

Theorem 3.4. *Let $r > 1$ and assume that the distribution μ of X is absolutely continuous with respect to λ having density f. Assume that $\operatorname{ess\,sup}_{\mathbb{R}} f < \infty$ and let $M = (\inf(\operatorname{supp}(\mu))), \sup(\operatorname{supp}(\mu)))$. In either of the following cases:*

(i) $\alpha \in (0,1)$, $E|X|^{r+\delta} < \infty$ *for some* $\delta > 0$*, and* f *is weakly unimodal,*

(ii) $\alpha \in (-\infty, 0)$, $\operatorname{ess\,inf}_M f > 0$ *and* f *is continuous on* M*,*

we have

$$\lim_{R\to\infty} e^{rR} D_\mu^\alpha(R) = C(r) \left(\int_M f^{a_1}\, d\lambda \right)^{a_2}. \tag{11}$$

If $\operatorname{ess\,inf}_M f > 0$ *and* f *is continuous on* M*, then*

$$\lim_{R\to\infty} e^{rR} D_\mu^{-\infty}(R) = C(r) \left(\int_M f^{1-r}\, d\lambda \right). \tag{12}$$

The proof of the theorem is given in Section 7. Upper bounds will be established using a companding approach, while matching lower bounds are developed by considering increasingly more general classes of source densities.

Remark 3.5. (a) Note that if we formally substitute $\alpha = 0$ in (11), it reduces to (8). Moreover, it is easy to show that (11) reduces to (9) if $\alpha \to 1$. Due to monotonicity of the quantization error (Lemma 2.3) and by the upper bound for the quantization error for $\alpha \in [-\infty, 1+r)$ (Corollary 4.11) one can rigorously show that the known asymptotics for $\alpha \in \{0,1\}$ also follow from Theorem 3.4, at least in the scalar case and under our restrictions on the source density.
(b) The results of the theorem can be expressed in terms of the Rényi differential entropy $h^\alpha(\mu) = \frac{1}{1-\alpha} \log\bigl(\int f^\alpha\, d\lambda\bigr)$ of order $\alpha \neq 1$. It is easy to check that (11) can be rewritten as

$$\lim_{R\to\infty} e^{rR} D_\mu^\alpha(R) = C(r) e^{r h^{a_1}(\mu)}. \tag{13}$$

Setting $a_1 = \lim_{\alpha\to-\infty} \frac{1-\alpha+\alpha r}{1-\alpha+r} = 1 - r$ for $\alpha = -\infty$, we also obtain (12) from the above expression. Also, for $\alpha = 0$ we have $a_1 = \frac{1}{1+r}$, and (13) reduces to (8); while for $\alpha = 1$, we have $a_1 = 1$, and

we formally get back (9) since $\lim_{a_1 \to 1} h^{a_1}(\mu) = h^1(\mu) = -\int f \log f \, d\lambda$ (cf. Section 4.2). Thus (13) expresses the old and the new asymptotic results in a unified form.

(c) Since $\operatorname{ess\,inf}_M f > 0$ the right hand side of (12) is finite. For the same reason, the right hand side of (11) is finite for all $\alpha < 0$. For $\alpha \in [0,1)$ the right hand side of (11) can be shown to be finite by an application of Hölder's inequality as in [9, Remark 6.3 (a)].

(d) The weak unimodality and continuity conditions on f are the results of our approximation techniques in proving lower bounds and are probably not necessary. In fact, with a a little tweaking of the companding approach in the next section one can show that the right hand sides of (11) and (12) still upper bound the asymptotic performance if these conditions are dropped.

(e) Note that condition (ii) implies (i). Also, the right hand side of (11) converges to the right hand side of (12) as $\alpha \to -\infty$.

(f) The condition $r > 1$ is needed in the proof of the lower bounds on $D_\mu^\alpha(R)$ where [19, Thm. 3.1] is invoked (see Proposition 7.2). The upper bounds only need $r \geq 1$ (see Section 4).

4 Distortion and Rényi entropy asymptotics of companding quantizers

4.1 Companding quantizers

Let $N \geq 2$ and $Q_N \in \mathcal{Q}$ denote the N-level uniform scalar quantizer with step size $1/N$ for sources supported in the unit interval $[0,1]$ defined by $Q_N(x) = 1/2N$ if $x \in [0, 1/N]$ and

$$Q_N(x) = \frac{i-1}{N} + \frac{1}{2N} \quad \text{if } x \in \left(\frac{i-1}{N}, \frac{i}{N} \right], \quad i = 2, \ldots, N. \tag{14}$$

The *compressor* G derived from a probability density g on the real line is the function

$$G(x) = \int_{-\infty}^{x} g(y) \, d\lambda(y). \tag{15}$$

Thus the increasing function $G : \mathbb{R} \to [0,1]$ is the cumulative distribution function associated with the density g. The generalized inverse $\hat{G}$ of G is defined by

$$\hat{G}(y) := \sup\{x : G(x) \leq y\} = \max\{x : G(x) \leq y\}$$

for $y \in (0,1)$. Note that if g is positive almost everywhere with respect to λ (a.e. for short), then G is strictly increasing and $\hat{G}$ is its (ordinary) inverse.

In this paper we will work only with compressor densities g having compact support, i.e., if ν denotes the measure induced by g, then $\operatorname{supp}(\nu)$ is bounded. Thus we can extend the definition of $\hat{G}$ onto $[0,1]$ by letting

$$\hat{G}(0) := \min\{\operatorname{supp}(\nu)\} > -\infty \quad \text{and} \quad \hat{G}(1) := \max\{\operatorname{supp}(\nu)\} < \infty.$$

The N-level *companding* quantizer $Q_{g,N}$ associated with g is defined on $[\hat{G}(0), \hat{G}(1)]$ by

$$Q_{g,N}(x) = \hat{G}(Q_N(G(x))).$$

Note that the codecells of $Q_{g,N}$ are N intervals $I_{1,N}, \ldots, I_{N,N}$ with $I_{1,N} = [\hat{G}(0), \hat{G}(1/N)]$ and

$$I_{i,N} = (\hat{G}((i-1)/N), \hat{G}(i/N)], \quad i = 2, \ldots, N.$$

The corresponding quantization points are $\hat{G}((2i-1)/2N), i = 1, \ldots, N.$

Remark 4.1. (a) The function g is often called the point density for $Q_{g,N}(x)$ since it has the property that for any $a < b$,

$$\lim_{N\to\infty} \frac{1}{N} \operatorname{card}(Q_{g,N}((a,b))) = \int_a^b g(x)\, d\lambda(x).$$

(b) If P_N is an arbitrary N-level quantizer on $\mathbb{R}$ having convex (interval) codecells, then it can be implemented as a companding quantizer. In particular, there exists a positive point density g such that $P_N(x) = Q_{g,N}(x)$ for all (except perhaps a finite number of) $x \in \mathbb{R}$ (any x such that $P_N(x) \neq Q_{g,N}(x)$ is a cell boundary for both quantizers).

The following result represents the error asymptotics of the compander if the number of output levels increases without bound. The result originates with Bennett [3] for $r = 2$ and has appeared in the literature in several different forms (but most often without precise conditions and a rigorous proof); see [12] for a historical overview. The proof is given in Appendix A and follows the development in [22] which gives a rigorous proof for the limit (16) under different conditions that include the continuity of g and certain tail conditions, but allow f and g to have unbounded support.

Proposition 4.2. *Let X be a random variable with distribution μ which is absolutely continuous with respect to λ and let f denote its density. Let G be a compressor with point density g. Assume that the support of μ is included in a compact interval I such that $\operatorname{ess\,inf}_I g > 0$ and $g(x) = 0$ a.e. on $\mathbb{R} \setminus I$. Then for $r \geq 1$,*

$$\lim_{N\to\infty} N^r D_\mu(Q_{g,N}) = C(r) \int_I \frac{f}{g^r}\, d\lambda. \tag{16}$$

Remark 4.3. Since $\operatorname{ess\,inf}_I g > 0$ we know that μ is absolutely continuous with respect to $g\lambda$ and $\int_I \frac{f}{g^r}\, d\lambda < \infty$.

4.2 Rényi entropy asymptotics of companding quantizers

In order to be able to construct asymptotically optimal companding quantizers (cf. (7)), in addition to the asymptotic distortion, we also have to control the quantizer's entropy, at least for high rates. In this section we derive a result (Proposition 4.8) which asymptotically describes the Rényi entropy of the compander as a function of the number of quantization points. Let 1_A denote the indicator function of $A \subset \mathbb{R}$.

Definition 4.4. *Let μ be absolutely continuous with respect to λ with density f and define $M = (\inf(\operatorname{supp}(\mu)), \sup(\operatorname{supp}(\mu)))$. Let $\alpha \in [-\infty, \infty]$ and assume that*

(i) $1_{\operatorname{supp}(\mu)} f^\alpha$ *is integrable if* $\alpha \in (-\infty, \infty) \setminus \{1\}$,

(ii) $\operatorname{ess\,inf}_M f > 0$ *if* $\alpha = -\infty$,

(iii) $f \log f$ *is integrable if* $\alpha = 1$,

(iv) $\operatorname{ess\,sup}_{\mathbb{R}} f < \infty$ *if* $\alpha = +\infty$.

Then the Rényi differential entropy of order α of μ is defined by

$$h^\alpha(\mu) = \begin{cases} \frac{1}{1-\alpha} \log(\int_{\operatorname{supp}(\mu)} f^\alpha \, d\lambda), & \alpha \in (-\infty, \infty) \setminus \{1\} \\ -\int f \log f \, d\lambda, & \alpha = 1 \\ -\log(\operatorname{ess\,sup}_{\mathbb{R}} f), & \alpha = \infty \\ -\log(\operatorname{ess\,inf}_M f), & \alpha = -\infty. \end{cases}$$

Remark 4.5. Just as in the case of Rényi entropy (see Remark 2.2) the mapping $[-\infty, \infty] \ni \alpha \to h^\alpha(\mu)$ is continuous for the differential entropy.

Recall that $\hat{q}_\Delta$ denotes the infinite-level uniform quantizer with step-size $\Delta > 0$. Recall M from Definition 4.4 and let $A(\Delta, M) = \{a \in \hat{q}_\Delta(\mathbb{R}) : \hat{q}_\Delta^{-1}(a) \subset M\}$ and

$$q_{\Delta,M}(\cdot) = \sum_{a \in A(\Delta,M)} a \cdot 1_{\hat{q}_\Delta^{-1}(a)}(\cdot).$$

The following result is due to Rényi [27] and Csiszár [8] for $\alpha \in (0, \infty)$. The proof for $\alpha \in [-\infty, 0]$ is given in Appendix A.

Lemma 4.6. *Let μ be absolutely continuous with respect to λ having density f. Let $\alpha \in [-\infty, \infty)$ and assume that the Rényi differential entropy of order α of μ exists and is finite. Assume that $H_\mu^\alpha(\hat{q}_\Delta) < \infty$ for some $\Delta > 0$. If $\alpha \in (-\infty, \infty)$, then*

$$\lim_{\Delta \to 0} \big(H_\mu^\alpha(\hat{q}_\Delta) + \log(\Delta) \big) = h^\alpha(\mu).$$

Moreover,

$$\lim_{\Delta \to 0} \big(H_\mu^{-\infty}(\hat{q}_{\Delta,M}) + \log(\Delta) \big) = h^{-\infty}(\mu).$$

Next we define the Rényi relative entropy between two probability measures for the case where both have densities.

Definition 4.7. *Let μ and ν be probability measures which are absolutely continuous with respect to λ. Denote by f and g the densities of μ and ν. Moreover, assume that μ is absolutely continuous with*

respect to ν and, therefore, we assume w.l.o.g. that $\{g=0\} \subset \{f=0\}$. Setting

$$E = \{f > 0\} \quad \textit{and} \quad M = (\inf(\operatorname{supp}(\mu)), \sup(\operatorname{supp}(\mu)))$$

the Rényi relative entropy of order α between the distributions μ and ν is defined as

$$\mathcal{D}_\alpha(\mu\|\nu) = \begin{cases} \frac{1}{\alpha-1}\log\Big(\int_E f^\alpha g^{1-\alpha}\, d\lambda\Big), & \alpha \in (-\infty,\infty)\setminus\{1\} \\ \int_E f \log\frac{f}{g}\, d\lambda, & \alpha = 1 \\ \log(\operatorname{ess\,sup}_E \frac{f}{g}), & \alpha = \infty \\ \log(\operatorname{ess\,inf}_M \frac{f}{g}), & \alpha = -\infty. \end{cases} \tag{17}$$

(For $\mathcal{D}_{-\infty}(\mu\|\nu)$ to be well defined, we need the condition $\operatorname{ess\,inf}_M f > 0$.)

The following result determines the asymptotics of the Rényi entropy of a companding quantizer.

Proposition 4.8. *Let $\alpha \in [-\infty,\infty)$. Suppose μ and ν are as in Definition 4.7 and $\mathcal{D}_\alpha(\mu\|\nu) < \infty$. Then*

$$\lim_{N\to\infty}\Big(H^\alpha_\mu(Q_{g,N}) - \log N\Big) = -\mathcal{D}_\alpha(\mu\|\nu).$$

Remark 4.9. (a) For the sake of distortion analysis we previously specified that g has bounded support, but in this proposition the only condition on f and g is the finiteness of $\mathcal{D}_\alpha(\mu\|\nu)$.
(b) In a sense, the proposition generalizes Lemma 4.6. Indeed, if the support of μ is included in a compact interval I and g is the uniform density on I, then $Q_{g,N}$ is the uniform quantizer of step-size $\Delta_N = \lambda(I)/N$ over I, and the proposition reduces to Lemma 4.6 (for the sequence of step-sizes Δ_N).

Proof. Recall the definition of the compressor G from (15). We proceed in two steps.

1. We show that $h^\alpha(\mu\circ G^{-1}) = -\mathcal{D}_\alpha(\mu\|\nu)$ for every $\alpha \in [-\infty,\infty)$.

Let $\alpha \in (-\infty,\infty)\setminus\{1\}$ and let f_G be the density of $\mu \circ G^{-1}$ (see Lemma A.1 in Appendix A). Definition 4.4 and Lemma A.1 imply

$$\begin{aligned} h^\alpha(\mu\circ G^{-1}) &= \frac{1}{1-\alpha}\log\int (f_G)^\alpha\, d\lambda \\ &= \frac{1}{1-\alpha}\log\int (f(\hat{G}(y))\hat{G}'(y))^\alpha\, d\lambda(y) \\ &= \frac{1}{1-\alpha}\log\Big(\int_{\hat{G}^{-1}(E)} f(\hat{G}(y))^\alpha g(\hat{G}(y))^{1-\alpha}\hat{G}'(y)\, d\lambda(y)\Big) \\ &= \frac{1}{1-\alpha}\log\Big(\int_E f(x)^\alpha g(x)^{1-\alpha}\, d\lambda(x)\Big) = -\mathcal{D}_\alpha(\mu\|\nu) \end{aligned}$$

where in the penultimate equality we used again the chain rule for the Lebesgue integral (see [30, Corollary 4]), which is applicable due to the monotonicity of $\hat{G}$ and the integrability of $f^\alpha g^{1-\alpha}$ (which follows

from the finiteness of $\mathcal{D}_\alpha(\mu\|\nu)$). Note that the above chain of equalities implies that $(f_G)^\alpha$ is integrable. One can deduce the assertion of step 1 for $\alpha \in \{-\infty, 1\}$ in a very similar manner.

2. Now we prove the assertion of the proposition. Since G is increasing and continuous, $\hat{G}$ is strictly increasing on $(0,1)$. Recall the definition of Q_N in (14) and note that $Q_N = \hat{q}_{1/N}$ on $(0,1)$. Then

$$H_\mu^\alpha(Q_{g,N}) = H_{\mu\circ G^{-1}}^\alpha(Q_N) \tag{18}$$

for all $\alpha \in [-\infty, \infty)$. Since $\mu \circ G^{-1}((0,1)) = 1$, we obtain $H_{\mu\circ G^{-1}}^\alpha(Q_N) = H_{\mu\circ G^{-1}}^\alpha(\hat{q}_{1/N})$. In view of (18) we deduce $H_\mu^\alpha(Q_{g,N}) = H_{\mu\circ G^{-1}}^\alpha(\hat{q}_{1/N})$. From step 1 and by the assumption we know that $h^\alpha(\mu \circ G^{-1})$ is finite. Since $\hat{q}_{1/N}$ has no more than N cells with nonzero $\mu \circ G^{-1}-$measure, the entropy $H_{\mu\circ G^{-1}}^\alpha(\hat{q}_{1/N})$ is also always finite. Lemma 4.6 and step 1 imply

$$\begin{aligned}\lim_{N\to\infty}\left(H_\mu^\alpha(Q_{g,N}) - \log N\right) &= \lim_{N\to\infty}\left(H_{\mu\circ G^{-1}}^\alpha(\hat{q}_{1/N}) - \log N\right)\\ &= h^\alpha(\mu\circ G^{-1}) = -\mathcal{D}_\alpha(\mu\|\nu).\end{aligned}$$

□

Remark 4.10. Although we do not need this fact in the sequel it is worth noting that Lemma 4.6 and Proposition 4.8 are also valid for $\alpha = \infty$. For example, by an application of Lebesgue's density theorem one can show that

$$\lim_{\Delta\to 0}\frac{\sup\{\mu(\hat{q}_\Delta^{-1}(a)) : a \in \hat{q}_\Delta(\mathbb{R})\}}{\Delta} = \operatorname{ess\,sup}_{\mathbb{R}} f,$$

which yields the assertion of Lemma 4.6 for $\alpha = \infty$. Generalizing the proof of Proposition 4.8 to $\alpha = \infty$ is straightforward.

4.3 Optimal point densities

Combining the previous results we can find a companding quantizer which provides an (asymptotic) upper bound for the optimal quantization error. Later on we will show that this quantizer is an asymptotically optimal one. Recall definition (10) of a_1 and a_2.

Corollary 4.11. *Let $r \geq 1$ and $\alpha \in [-\infty, 1+r)$. Assume that μ is supported on a compact interval I and has density f such that $\operatorname{ess\,inf}_I f > 0$. Moreover, assume that f^{a_1} is integrable if $\alpha \in (1, 1+r)$ and $f \log f$ is integrable if $\alpha = 1$. Let*

$$f^* = \begin{cases} (\int_I f^{1/a_2}\, d\lambda)^{-1} f^{1/a_2}, & \alpha \in (-\infty, 1+r)\setminus\{1\}\\ (\lambda(I))^{-1} 1_I, & \alpha = 1\\ f, & \alpha = -\infty.\end{cases} \tag{19}$$

Then,

$$\lim_{N\to\infty} e^{rH_\mu^\alpha(Q_{f^*,N})} D_\mu(Q_{f^*,N}) = \begin{cases} C(r)(\int_I f^{a_1}\, d\lambda)^{a_2}, & \alpha \in (-\infty, 1+r) \setminus \{1\} \\ C(r) e^{-r\int f \log f\, d\lambda}, & \alpha = 1 \\ C(r) \int_I f^{1-r}\, d\lambda, & \alpha = -\infty. \end{cases} \tag{20}$$

Proof. It is not hard to show using Hölder's inequality that $f^{a_1} 1_I$ is integrable for every $\alpha \in (-\infty, 1+r) \setminus \{1\}$ (cf. [9, Remark 6.3 (a)]). Moreover f^{1-r} is integrable. Clearly, $f^{1/a_2} 1_I$ is integrable for every $\alpha \in (-\infty, 1+r)$. These facts imply that f^* is well defined (note that $\operatorname{ess\,inf}_I f^* > 0$), $\int f/(f^*)^r\, d\lambda < \infty$, and the integrals on the right hand side of (20) are finite. It is also easy to check that $\mathcal{D}_\alpha(\mu \| f^*\lambda)$ is finite. Thus we can apply Propositions 4.8 and 4.2. We obtain

$$\begin{aligned} \lim_{N\to\infty} e^{rH_\mu^\alpha(Q_{f^*,N})} D_\mu(Q_{f^*,N}) &= \lim_{N\to\infty} e^{-r\mathcal{D}_\alpha(\mu\|f^*\lambda)} N^r D_\mu(Q_{f^*,N}) \\ &= e^{-r\mathcal{D}_\alpha(\mu\|f^*\lambda)} C(r) \int_I \frac{f}{(f^*)^r}\, d\lambda. \end{aligned}$$

Now (19) and (17) yield the assertion. □

Remark 4.12. For $\alpha \in [-\infty, 1+r)$ the point density $g = f^*$ in the corollary minimizes the asymptotic performance $\lim_{N\to\infty} e^{rH_\mu^\alpha(Q_{g,N})} D_\mu(Q_{g,N})$. For $\alpha = 0$ and $\alpha = 1$ this optimal choice of g has long been known. In the case $\alpha \in (-\infty, 1+r) \setminus \{1\}$, by Propositions 4.2 and 4.8 the above limit is proportional to

$$\left(\int_I f^\alpha g^{1-\alpha}\, d\lambda \right)^{\frac{r}{1-\alpha}} \int_I f g^{-r}\, d\lambda$$

and Hölder's inequality (for $\alpha < 1$) or the reverse Hölder inequality (for $\alpha \in (1, 1+r)$) can be used to show that this functional is minimized by $g = f^*$. The resulting minimum is $\left(\int_I f^{a_1}\, d\lambda\right)^{a_2}$. The case $\alpha = -\infty$ follows by letting $\alpha \to -\infty$.

5 Some important properties of optimal scalar quantization

Define

$$i(f) = \operatorname{ess\,inf}_{\operatorname{supp}(\mu)} f, \qquad s(f) = \operatorname{ess\,sup}_{\operatorname{supp}(\mu)} f.$$

For the case $\alpha = 0$ the following result is originally due to Pierce ([25], [9, Lemma 6.6]). In our proof, given in Appendix B, we use a refined version provided by Luschgy and Pagès [23, Lemma 1].

Proposition 5.1. *(i) If $R \geq 1$ and $\int |x|^{r+\beta}\, d\mu(x) < \infty$ for some $\beta > 0$, then there exists a constant $C_0 > 0$ (which depends only on r and β) such that*

$$e^{rR} D_\mu^\alpha(R) \leq C_0 \left(\int |x|^{r+\beta}\, d\mu(x) \right)^{r/(r+\beta)}$$

for every $\alpha \in [0,\infty]$.
(ii) Suppose $\mathrm{supp}(\mu)$ *is a compact interval and* μ *absolutely continuous with respect to* λ *with density* f. *Assume that* $i(f) > 0$. *Then for all* $\alpha < 0$

$$e^{rR} D_\mu^\alpha(R) \le \frac{2^r}{i(f)^r}. \tag{21}$$

As an immediate consequence we obtain the following.

Corollary 5.2. *Under either condition (i) or (ii) of Proposition 5.1 we have* $\lim_{R\to\infty} D_\mu^\alpha(R) = 0$.

Let $\mathrm{diam}(A) = \sup\{|x-y| : x,y \in A\}$ denote the diameter of an arbitrary non-empty set $A \subset \mathbb{R}$. The next result shows that the measure of the codecells of optimal quantizers tends to zero for absolutely continuous distributions. The proof, given in Appendix B, adopts some techniques of Gray *et al*. [11, Proof of Lemma 11].

Lemma 5.3. *Let* μ *be absolutely continuous with respect to* λ *having density* f. *Assume further either of the following conditions*

(i) $\alpha \in [0,\infty]$ *and* $\int |x|^{r+\beta}\, d\mu(x) < \infty$ *for some* $\beta > 0$,

(ii) $\alpha < 0$ *and* $\mathrm{supp}(\mu)$ *is a compact interval and* $0 < i(f) \le s(f) < \infty$.

Then for every $\varepsilon > 0$ *there exists an* $R_0 > 0$ *with the property that for every* $R \ge R_0$ *there is a* $\delta > 0$ *such that*

$$\max\{\mu(q^{-1}(a)) : a \in q(\mathbb{R})\} < \varepsilon \tag{22}$$

for every $q \in \mathcal{Q}$ *with* $H_\mu^\alpha(q) \le R$ *and* $|D_\mu(q) - D_\mu^\alpha(R)| < \delta$. *If, additionally, in case (i) the support of* μ *consists of* $m \ge 1$ *compact intervals* $I_1,\dots,I_m$ *and* $i(f) > 0$, *then in both cases (i) and (ii) we have*

$$\max\{\mathrm{diam}(q^{-1}(a) \cap I_i) : a \in q(\mathbb{R}), i \in \{1,\dots,m\}\} < \varepsilon \cdot i(f)^{-1} \tag{23}$$

where $m = 1$ *and* $I_1 = I$ *for case (ii).*

Let μ be absolutely continuous with respect to λ and denote the density of μ with f. Let

$$C = \left(\frac{i(f)}{s(f)}\right)^{\frac{r+1}{r}} \left(\frac{1}{4^r(1+r)}\right)^{1/r} \in (0,1). \tag{24}$$

For any $q \in \mathcal{Q}$ let

$$N_q = \{a \in q(\mathbb{R}) : \mu(q^{-1}(a)) > 0\}.$$

In the case $\alpha < 0$ we need to control in our proofs the cardinality of the codebook of any quantizer whose entropy is less than or equal to the rate constraint R. To this end, for $R \ge 0$, we define

$$\mathcal{H}_R = \{q \in \mathcal{Q}^c : \quad H_\mu^\alpha(q) \le R, \quad Ce^R \le \mathrm{card}(N_q) \le e^R\}.$$

In addition, we will have to control the difference between the rate constraint and the entropy of the quantizer. Thus, for $\alpha \in (-\infty, 0)$, arbitrary constant $\kappa > 0$, and $R > \log(\frac{2^{1-\alpha}-1}{\kappa})$, we define

$$\mathcal{K}_R = \mathcal{K}_R(\kappa) = \left\{ q \in \mathcal{H}_R : e^{R - H_\mu^\alpha(q)} \leq \left(\frac{1}{1 - (2^{1-\alpha} - 1)\kappa^{-1} e^{-R}} \right)^{1/(1-\alpha)} \right\}.$$

The next lemma is proved in Appendix B.

Lemma 5.4. *Let μ be absolutely continuous with respect to λ having density f. Assume that $\operatorname{supp}(\mu)$ is a compact interval and $0 < i(f) \leq s(f) < \infty$. For every $\alpha \in [-\infty, 0]$ and $R \geq 0$ we have*

$$D_\mu^\alpha(R) = \inf\{D_\mu(q) : q \in \mathcal{H}_R\}. \tag{25}$$

If $\alpha \in (-\infty, 0)$ and $R > \log(\frac{2^{1-\alpha}-1}{C})$, then

$$D_\mu^\alpha(R) = \inf\{D_\mu(q) : q \in \mathcal{K}_R(C)\}. \tag{26}$$

We let $U(I)$ denote the uniform distribution on a bounded interval $I \subset \mathbb{R}$ with positive length. Let $m \geq 2$ and let $I_1, \ldots, I_m$ be a partition of I into m intervals of equal length $\operatorname{diam}(I)/m$. Let $s_1, \ldots, s_m \in (0,1)^m$ with $\sum_{i=1}^m s_i = 1$ and assume the source distribution is of the form $\mu = \sum_{i=1}^m s_i U(I_i)$. Of special interest in our proofs are the codecells which are straddling the intervals I_i. Hence we define for any quantizer $q \in \mathcal{Q}$ the sets

$$A(q) = \bigcup_{i=1}^m \{a \in q(\mathbb{R}) : \lambda(q^{-1}(a) \setminus I_i) = 0\} \quad \text{and} \quad S(q) = q(\mathbb{R}) \setminus A(q). \tag{27}$$

In the proof of our main result we have to ensure that the contribution of the straddling cells to the overall entropy of the quantizer can be (asymptotically) neglected. For $\alpha < 0$ this is the case if it suffices to consider only quantizers with the property that the length of each straddling cell is at least as large as a certain (fixed) constant times the length of the smallest non-straddling cell. Exactly this is ensured by the following lemma which sharpens Lemma 5.4. The proof is given in Appendix B. Recall the definition (24) of the constant C and let $\kappa \in (0, C)$. For $R > \log(\frac{2^{1-\alpha}-1}{\kappa})$ let

$$\begin{aligned} \mathcal{G}_R = \mathcal{G}_R(\kappa) \quad &= \quad \big\{ q \in \mathcal{K}_R(\kappa) : 2\inf\{\operatorname{diam}(q^{-1}(a) \cap I) : a \in S(q)\} \\ &\qquad \geq \inf\{\operatorname{diam}(q^{-1}(a)) : a \in A(q)\} \big\}. \end{aligned}$$

Lemma 5.5. *Assume that $\mu = \sum_{i=1}^m s_i U(I_i)$ is a piecewise uniform distribution as specified above. Let $r > 1$ and $\kappa \in (0, C)$. Then for every $\alpha \in (-\infty, 0)$ there is an $R_0(\kappa) > 0$ such that for every $R \geq R_0(\kappa)$,*

$$D_\mu^\alpha(R) = \inf\{D_\mu(q) : q \in \mathcal{G}_R(\kappa)\}.$$

A bijective mapping $T : \mathbb{R} \to \mathbb{R}$ is called a similarity transformation if there exists $c \in (0, \infty)$,

the scaling number, such that $|Tx - Ty| = c|x - y|$ for every $x, y \in \mathbb{R}$. The last result of this section describes how the optimal quantization error scales under a similarity transformation. For $\alpha = 0$ the reader is also referred to [9, Lemma 3.2]. Let us denote by

$$C_\mu^\alpha(R) = \{q \in \mathcal{Q}^c : D_\mu(q) = D_\mu^\alpha(R)\}$$

the set of all optimal quantizers in $\mathcal{Q}^c$ for μ under Rényi entropy constraint R of order α.

Lemma 5.6. *Let $\alpha \in [-\infty, \infty]$ and $T : \mathbb{R} \to \mathbb{R}$ be a similarity transformation with scaling number $c > 0$. Then for any $R \geq 0$ we have*

$$D_{\mu \circ T^{-1}}^\alpha(R) = c^r D_\mu^\alpha(R).$$

Moreover,

$$C_{\mu \circ T^{-1}}^\alpha(R) = \{T \circ q \circ T^{-1} : q \in C_\mu^\alpha(R)\}.$$

Proof. The lemma follows because for any $q \in \mathcal{Q}$ we have $\bar{q} := T \circ q \circ T^{-1} \in \mathcal{Q}$, $H_{\mu \circ T^{-1}}^\alpha(q) = H_\mu^\alpha(\bar{q})$, and $D_{\mu \circ T^{-1}}(q) = c^r D_\mu(\bar{q})$ (also, $q \in \mathcal{Q}^c$ iff $\bar{q} \in \mathcal{Q}^c$). See also [19, Lemma 2.4] where $\alpha \geq 0$ and $r > 1$ are considered, but the same proof clearly works for all $\alpha < 0$ and $r > 0$. □

6 Inequalities for mixture distributions

In this section we provide upper and lower bounds for the optimal quantization error of mixture distributions in terms of the optimal quantizer performance for the component distributions. Proofs are given in Appendix C.

Definition 6.1. *Let $m \geq 2$ and $A_1, \ldots, A_m$ be measurable sets which are pairwise disjoint. The distribution μ is called $m-$divisible with respect to $(A_1, \ldots, A_m)$ if $\mu(A_i) > 0$ for all $i = 1, \ldots, m$ and $\mu(\cup_{i=1}^m A_i) = 1$.*

For any measurable $A \subset \mathbb{R}$ with $\mu(A) > 0$ we let $\mu(\cdot|A)$ denote the conditional probability of μ with respect to A, i.e., $\mu(B|A) = \mu(B \cap A)/\mu(A)$ for all measurable $B \subset \mathbb{R}$. If μ is $m-$divisible, then we write $\mu_i = \mu(\cdot|A_i)$.

Proposition 6.2. *Let $R \geq 0$, $\alpha \in [0, \infty) \setminus \{1\}$, and $m \geq 2$. Assume that μ is $m-$divisible with partition $(A_1, \ldots, A_m)$. Moreover assume, that $\int |x|^r d\mu_i(x) < \infty$ for every $i = 1, \ldots, m$. Let $R_1, \ldots, R_m \in [0, \infty)$. Letting $s_i = \mu(A_i)$, we have*

$$D_\mu^\alpha(R) \leq \sum_{i=1}^m s_i D_{\mu_i}^\alpha(R_i)$$

if either one of the following inequalities holds:

$$\log\left(\sum_{i=1}^{m} s_i^{\alpha} e^{(1-\alpha)R_i}\right) \leq (1-\alpha)R \qquad \textit{if } \alpha \in [0,1), \tag{28}$$

$$\log\left(\sum_{i=1}^{m} s_i^{\alpha} e^{(1-\alpha)R_i}\right) \geq (1-\alpha)R \qquad \textit{if } \alpha \in (1,\infty). \tag{29}$$

Recall the definition (10) of a_1 and a_2. Let $m \geq 2$ and $s_1, \ldots, s_m \in (0,1)^m$ with $\sum_{i=1}^m s_i = 1$. For every $i \in \{1, \ldots, m\}$ and $\alpha \in [0, r+1) \setminus \{1\}$ let

$$t_i = s_i^{1/a_2} \left(\sum_{j=1}^{m} s_j^{a_1}\right)^{-\frac{1}{1-\alpha}}. \tag{30}$$

Lemma 6.3. *Let $m \geq 2$. Let μ be non-atomic and $m-$divisible with respect to $(A_1, \ldots, A_m)$. Assume $\int |x|^r \, d\mu_i(x) < \infty$ for all $i = 1, \ldots, m$. Let $i_0 \in \{1, \ldots, m\}$ with $\mu(A_{i_0}) = s = \max\{\mu(A_i) : i = 1, \ldots, m\}$. If $\alpha \in [0,1)$, then*

$$\liminf_{R\to\infty} e^{rR} D_{\mu}^{\alpha}(R) \geq s^{a_1 a_2} \liminf_{R\to\infty} e^{rR} D_{\mu_{i_0}}^{\alpha}(R). \tag{31}$$

Let $s_i = \mu(A_i)$ and assume $\alpha \in [0, r+1) \setminus \{1\}$. Then we have

$$\limsup_{R\to\infty} e^{rR} D_{\mu}^{\alpha}(R) \leq \sum_{i=1}^{m} s_i t_i^{-r} \limsup_{R\to\infty} e^{rR} D_{\mu_i}^{\alpha}(R). \tag{32}$$

7 Proof of main result

Recall that $U(I)$ denotes the uniform distribution on a bounded interval I with positive length. First we show that the optimal quantizer performance for $U(I)$ is the same for all negative α.

Lemma 7.1. *Let $-\infty < a < b < \infty$. For every $R \geq 0$ and $\alpha < 0$, we have*

$$D_{U([a,b])}^{\alpha}(R) = D_{U([a,b])}^{0}(R).$$

Proof. Note that by Lemma 5.6 it suffices to consider the case $[a,b] = [0,1]$. Since $D_{U([a,b])}^{\alpha}(R)$ is nonincreasing in α by Lemma 2.3 it suffices to prove the assertion for $\alpha = -\infty$. Let $R \geq 0$ and assume $q \in \mathcal{Q}$ satisfies $H_{\mu}^{-\infty}(q) \leq R$. Setting

$$N_q = \{a \in q(\mathbb{R}) : \mu(q^{-1}(a)) > 0\}$$

this condition is equivalent to

$$p = \min\{\mu(q^{-1}(a)) : a \in N_q\} \geq \exp(-R). \tag{33}$$

Let $\lfloor x \rfloor$ denote the largest integer less than or equal to $x \in \mathbb{R}$. Using $1 \geq \mathrm{card}(N_q) \cdot p$ we get

$$\mathrm{card}(N_q) \leq \lfloor \exp(R) \rfloor, \tag{34}$$

which is equivalent to $H^0_\mu(q) \leq R$. From, e.g., [9, Example 5.5] we know that only the quantizer $g \in \mathcal{Q}$ which partitions the unit interval into $\lfloor \exp(R) \rfloor$ intervals of equal length with their midpoints as quantization points, attains the optimal error, i.e., $D_{U([0,1])}(g) = D^0_{U([0,1])}(R)$. But this quantizer satisfies conditions (33) and (34) simultaneously. Hence, $D_{U([0,1])}(g) = D^{-\infty}_{U([0,1])}(R)$, which yields the assertion. $\square$

Next we determine the exact behavior of $D^\alpha_{U([0,1])}(R)$ for large R. For $\alpha = 1$ the following result is from [13]. For the case $\alpha = 0$ the reader is referred, for example, to [9, Example 5.5].

Proposition 7.2. *Let $r > 1$ and $R > 0$. Let $-\infty < a < b < \infty$. Then the following hold:*

(i) If $\alpha \in [0, r+1)$, an optimal quantizer always exists for $U([a,b])$, i.e., we can find a $q \in \mathcal{Q}$ with $H^\alpha_{U([a,b])}(q) \leq R$ and $D_{U([a,b])}(q) = D^\alpha_{U([a,b])}(R)$.

(ii) Suppose $\alpha \in [0, r+1)$ and let $n \in \mathbb{N}$ be such that $R \in (\log(n), \log(n+1)]$. Then the restriction to $[a,b]$ of the quantizer q in (i) has $(n+1)$ interval cells, n of which are of equal lengths and one having length less than or equal to that of the others. If $\alpha > 0$, then q meets the entropy constraint with equality, i.e., $H^\alpha_{U([a,b])}(q) = R$.

(iii) For all $\alpha \in [-\infty, r+1)$, we have

$$\lim_{R\to\infty} e^{rR} D^\alpha_{U([0,1])}(R) = C(r). \tag{35}$$

Proof. Assertions (i) and (ii) follow directly from [19, Thm. 3.1] by noting that in view of Lemma 5.6 it suffices to consider the case $[a,b] = [0,1]$

To prove (iii) first we note that by Lemma 7.1, the limit (35) holds for all $\alpha \in [-\infty, 0)$ since it holds for $\alpha = 0$. Thus we need only concentrate on the case $\alpha \in (0, r+1) \setminus \{1\}$. Applying [19, Thm. 3.1] we obtain

$$D^\alpha_{U([0,1])}(\log(n)) = C(r) n^{-r}$$

for every $n \in \mathbb{N}$. Now let $R \geq 0$ and $n_R \in \mathbb{N}$, such that $\log(n_R) < R \leq \log(n_R + 1)$. We get

$$\begin{aligned} n_R^r C(r)(n_R+1)^{-r} &= n_R^r D^\alpha_{U([0,1])}(\log(n_R+1)) \\ &\leq e^{rR} D^\alpha_{U([0,1])}(R) \leq (n_R+1)^r D^\alpha_{U([0,1])}(\log n_R) = (n_R+1)^r C(r) n_R^{-r}. \end{aligned}$$

Letting $R \to \infty$ yields (35) for $\alpha \in (0, r+1) \setminus \{1\}$. $\square$

Proof of Theorem 3.4. We divide the proof into four main steps. In step 1 we begin by proving a (sharp) asymptotic lower bound on the optimal quantization error for any distribution with a density that is piecewise constant on a finite number of intervals of equal lengths. In step 2 we generalize the lower bound of step 1 to any density whose support is a compact interval on which it is bounded away from zero. Together with a matching upper bound based on the companding result Corollary 4.11 this will finish the proof for $\alpha \in (-\infty, 0)$. In step 3 we show that the lower bound holds for all distributions subject to our restrictions and apply again the companding upper bound to finish the proof for $\alpha \in (0, 1)$. Step 4 treats the remaining $\alpha = -\infty$ case and thus completes the proof.

Throughout we assume w.l.o.g. that $R \geq R_0(C/2)$ where is C defined in (24) and $R_0(C/2)$ is from Lemma 5.5.

Step 1.

Let M be a compact interval of positive length and let $m \geq 2$ and $\alpha \in (-\infty, 1) \setminus \{0\}$. Assume that $\mu = \sum_{i=1}^m s_i U(A_i)$, where the A_i are disjoint intervals of equal length $l = l(A_i) = \lambda(M)/m$ that form a partition of M. We assume $s_i > 0$ for all $i = 1, \dots, m$. Thus $\sum_{i=1}^m s_i = 1$ and

$$f = \frac{d\mu}{d\lambda} = \sum_{i=1}^m s_i l^{-1} 1_{A_i}.$$

For $\alpha \in (-\infty, 1) \setminus \{0\}$ define $t_i = s_i^{1/a_2} \left(\sum_{j=1}^m s_j^{a_1} \right)^{-\frac{1}{1-\alpha}}$, $i = 1, \dots, m$ as in (30). Let

$$R \geq \max\{0, \max\{-\log(t_i) : i = 1, \dots, m\}\}$$

and define

$$R_i = R + \log(t_i) \geq 0.$$

From Proposition 7.2 we deduce

$$\begin{aligned} e^{rR} D^\alpha_{U([0,1])}(R_i) &= \left(e^{R-R_i}\right)^r e^{rR_i} D^\alpha_{U([0,1])}(R_i) \\ &\to t_i^{-r} C(r) \qquad \text{as } R \to \infty. \end{aligned}$$

A simple calculation shows

$$\begin{aligned} \left(\int_M f^{a_1} \, d\lambda \right)^{a_2} &= \left(\int \left(\sum_{i=1}^m s_i l^{-1} 1_{A_i} \right)^{a_1} d\lambda \right)^{a_2} \\ &= l^{(1-a_1)a_2} \left(\sum_{i=1}^m s_i^{a_1} \right)^{a_2} \\ &= l^r \left(\sum_{i=1}^m s_i^{a_1} \right)^{a_2} = l^r \sum_{i=1}^m s_i t_i^{-r}. \end{aligned} \tag{36}$$

Now, according to Lemma 5.3 there exist functions $\varepsilon : (0,\infty) \to (0,\infty)$ and $\delta : (0,\infty) \to (0,\infty)$ such that for every $R > 0$ and quantizer $q \in \mathcal{Q}$ with $H_\mu^\alpha(q) \le R$ and $|D_\mu^\alpha(R) - D_\mu(q)| \le \delta(R)$ we have

$$\max\{\mu(q^{-1}(a)) : a \in q(\mathbb{R})\} < \varepsilon(R) \tag{37}$$

where $\varepsilon(R) \to 0$ as $R \to \infty$. Moreover,

$$\max\{\mathrm{diam}(A_i \cap q^{-1}(a)) : a \in q(\mathbb{R}), i \in \{1, \ldots, m\}\} < \frac{l \cdot \varepsilon(R)}{\min\{s_i : i = 1, \ldots, m\}}. \tag{38}$$

Now, again, let $R \ge R_0$ and $\gamma > 0$. According to Lemma 2.3 let $q_R \in \mathcal{Q}$ be a quantizer whose codecells with positive $\mu-$mass are intervals, satisfying $H_\mu^\alpha(q_R) \le R$ and

$$|D_\mu^\alpha(R) - D_\mu(q_R)| \le \min(\gamma e^{-rR}, \delta(R)). \tag{39}$$

Hence, q_R satisfies also the relations (37) and (38). In view of Lemma 5.5 let us assume w.l.o.g. that $q_R \in \mathcal{G}_R$ if $\alpha < 0$. Now let $i \in \{1, \ldots, m\}$ and

$$I_i(q_R) = \{a \in q_R(\mathbb{R}) : \lambda\big(q_R^{-1}(a) \setminus A_i\big) = 0\}$$

and

$$A_{i,q_R} = \bigcup_{a \in I_i(q_R)} q_R^{-1}(a).$$

With

$$J_i(q_R) = \{a \in q_R(\mathbb{R}) \setminus I_i(q_R) : \mu(A_i \cap q_R^{-1}(a)) > 0\}$$

we obtain from (38) that

$$\lim_{R\to\infty} \sup\{\mathrm{diam}(A_i \cap q_R^{-1}(a)) : a \in J_i(q_R)\} = 0.$$

Every point of $J_i(q_R)$ is a codepoints of a codecell which is straddling the boundary of A_i and is not $\mu-$a.s. contained in A_i. Hence $\{A_i \cap q_R^{-1}(a) : a \in J_i(q_R)\}$ consists of at most two intervals and we get

$$\lim_{R\to\infty} \mathrm{diam}(A_{i,q_R}) = \mathrm{diam}(A_i) = l. \tag{40}$$

We compute

$$\begin{aligned} D_\mu(q_R) &= \sum_{i=1}^m s_i l^{-1} \int_{A_i} |x - q_R(x)|^r \, d\lambda(x) \\ &\ge \sum_{i=1}^m s_i l^{-1} \int_{A_{i,q_R}} |x - q_R(x)|^r \, d\lambda(x). \end{aligned} \tag{41}$$

Let

$$
\begin{aligned}
R_{i,q_R} &= H^{\alpha}_{U(A_{i,q_R})}(q_R) \\
&= \frac{1}{1-\alpha}\log\left(\sum_{a\in I_i(q_R)}(U(A_{i,q_R})(q_R^{-1}(a)))^{\alpha}\right) \\
&= \frac{\alpha}{\alpha-1}\left(\log(s_i)-\log\left(\frac{l}{\lambda(A_{i,q_R})}\right)\right)+\frac{1}{1-\alpha}\log\left(\sum_{a\in I_i(q_R)}\mu(q_R^{-1}(a)))^{\alpha}\right) \qquad (42)
\end{aligned}
$$

where $U(A_{i,q_R})(q_R^{-1}(a))$ is the measure of the cell $q_R^{-1}(a)$ under the uniform distribution on A_{i,q_R}. Then using Lemma 5.6 and (41) we obtain

$$
\begin{aligned}
D_\mu(q_R) &\geq \sum_{i=1}^{m} s_i l^{-1} D^{\alpha}_{U(A_{i,q_R})}(R_{i,q_R})\,\mathrm{diam}(A_{i,q_R}) \\
&= \sum_{i=1}^{m} s_i l^{-1} D^{\alpha}_{U([0,1])}(R_{i,q_R})\,\mathrm{diam}(A_{i,q_R})^{1+r}. \qquad (43)
\end{aligned}
$$

Now pick a sequence (L_n) of non-negative real numbers, such that $L_n \to \infty$,

$$
e^{rL_n} D^{\alpha}_{\mu}(L_n) \to \liminf_{R\to\infty} e^{rR} D^{\alpha}_{\mu}(R), \qquad (44)
$$

and

$$
\frac{e^{R_{i,q_{L_n}}}}{e^{L_n}} \to v_i \in [0,\infty], \qquad i = 1,\ldots,m \qquad (45)
$$

as $n \to \infty$. Because we want to determine a lower bound for the optimal quantization error, using Proposition 7.2 (i) we can assume w.l.o.g. that q_{L_n} is $R_{i,q_{L_n}}-$optimal for $U(A_{i,q_{L_n}})$, i.e., that $D^{\alpha}_{U(A_{i,q_{L_n}})}(R_{i,q_{L_n}}) = D_{U(A_{i,q_{L_n}})}(q_{L_n})$. By Proposition 7.2 (ii) the quantizer q_{L_n} divides $A_{i,q_{L_n}}$ into $(k+1)$-intervals with $R_{i,q_{L_n}} \in (\log(k), \log(k+1)]$ where at least k intervals are of equal length.

We next prove that $R_{i,q_{L_n}} \to \infty$ as $n \to \infty$. Assume to the contrary that $(R_{i,q_{L_n}})_{n\in\mathbb{N}}$ is bounded. Then $k = k(R_{i,q_{L_n}})$ will also be bounded. Thus let $k_0 \in \mathbb{N}$ such that $k \in \{1,\ldots,k_0\}$ for every $n \in \mathbb{N}$. Together with (40) we deduce

$$
\begin{aligned}
\liminf_{n\to\infty} D_\mu(q_{L_n}) &\geq \liminf_{n\to\infty} \int_{A_{i,q_{L_n}}} |x - q_{L_n}(x)|^r \, d\mu(x) \\
&\geq \liminf_{n\to\infty} \frac{s_i}{l} k \cdot 2 \int_0^{\frac{1}{2}\frac{\mathrm{diam}(A_{i,q_{L_n}})}{k+1}} x^r \, d\lambda(x) \\
&\geq C(r) s_i l^r \min\left\{\frac{k}{(k+1)^{r+1}} : k \in \{1,\ldots,k_0\}\right\} > 0.
\end{aligned}
$$

But this contradicts (cf. Corollary 5.2)

$$\limsup_{n\to\infty} D_\mu(q_{L_n}) \le \limsup_{n\to\infty}(D_\mu^\alpha(L_n)+\gamma e^{-rL_n}) = 0.$$

Thus we obtain that $R_{i,q_{L_n}} \to \infty$ as $n\to\infty$ for all $i\in\{1,\ldots,m\}$. Proposition 7.2 yields

$$\lim_{n\to\infty} e^{rR_{i,q_{L_n}}} D^\alpha_{U([0,1])}(R_{i,q_{L_n}}) = C(r), \qquad i=1,\ldots,m. \tag{46}$$

Because $\gamma>0$ was arbitrary we obtain

$$\begin{aligned}
\liminf_{R\to\infty} e^{rR} D_\mu^\alpha(R) = \lim_{n\to\infty} e^{rL_n} D_\mu^\alpha(L_n) &= \lim_{n\to\infty} e^{rL_n} D_\mu(q_{L_n}) \\
&\ge \lim_{n\to\infty} e^{L_n r}\sum_{i=1}^m s_i l^{-1} D^\alpha_{U([0,1])}(R_{i,q_{L_n}})\,\mathrm{diam}(A_{i,q_{L_n}})^{1+r} \\
&= C(r)\sum_{i=1}^m s_i v_i^{-r} l^r
\end{aligned} \tag{47}$$

where the first equality holds by (44), the second by (39), the inequality follows from (43), and the third equality follows from (40), (45), and (46). In the last expression, $1/v_i=0$ if $v_i=\infty$. The case $v_i=0$ cannot occur because otherwise the right hand side of (47) is not finite, which would contradict the assertion of Proposition 5.1. Recall that $\{A_i\cap q_R^{-1}(a) : a\in J_i(q_R)\}$ contains at most two intervals for every i and $n\in\mathbb{N}$. Now assume that $\alpha\in(0,1)$. In this case, since by (6) we can assume w.l.o.g. that $H_\mu^\alpha(q_{L_n}) = L_n$, we obtain

$$\begin{aligned}
1 \;&\le\; \delta_1(L_n,\mu,q_{L_n}) := \frac{e^{L_n(1-\alpha)}}{\sum_{i=1}^m\sum_{a\in I_i(q_{L_n})}\mu(q_{L_n}^{-1}(a))^\alpha} \\
&= \frac{\sum_{i=1}^m\sum_{a\in I_i(q_{L_n})}\mu(q_{L_n}^{-1}(a))^\alpha + \sum_{a\in q_{L_n}(\mathbb{R})\setminus\cup_{j=1}^n I_j(q_{L_n})}\mu(q_{L_n}^{-1}(a))^\alpha}{\sum_{i=1}^m\sum_{a\in I_i(q_{L_n})}\mu(q_{L_n}^{-1}(a))^\alpha} \\
&\le \frac{\sum_{i=1}^m\sum_{a\in I_i(q_{L_n})}\mu(q_{L_n}^{-1}(a))^\alpha + (m+1)\sup_{a\in q_{L_n}(\mathbb{R})}\mu(q_{L_n}^{-1}(a))^\alpha}{\sum_{i=1}^m\sum_{a\in I_i(q_{L_n})}\mu(q_{L_n}^{-1}(a))^\alpha} \\
&= 1+\frac{(m+1)\sup_{a\in q_{L_n}(\mathbb{R})}\mu(q_{L_n}^{-1}(a))^\alpha}{\sum_{i=1}^m\sum_{a\in I_i(q_{L_n})}\mu(q_{L_n}^{-1}(a))^\alpha}.
\end{aligned}$$

From (42) and $\lim_{n\to\infty} R_{i,q_{L_n}}=\infty$ we deduce

$$\lim_{n\to\infty}\sum_{i=1}^m\sum_{a\in I_i(q_{L_n})}\mu(q_{L_n}^{-1}(a))^\alpha = \infty.$$

Thus we get

$$\lim_{n\to\infty}\delta_1(L_n,\mu,q_{L_n}) = 1. \tag{48}$$

Using Lemma D.1 in Appendix D we recognize that the limit relation (48) also holds for $\alpha < 0$. Consequently, we deduce together with (40) and (42) for every $\alpha \in (-\infty, 1) \setminus \{0\}$ that

$$\begin{aligned}\sum_{i=1}^{m} s_i^{\alpha} v_i^{1-\alpha} &= \sum_{i=1}^{m} s_i^{\alpha} \lim_{n\to\infty} e^{(R_{i,q_{L_n}} - L_n)(1-\alpha)} \\ &= \lim_{n\to\infty} \left(\sum_{i=1}^{m} \sum_{a \in I_i(q_{L_n})} \mu(q_{L_n}^{-1}(a))^{\alpha} \left(\frac{l}{\lambda(A_{i,q_{L_n}})} \right)^{\alpha} \right) e^{-L_n(1-\alpha)} = 1.\end{aligned}$$

Moreover we obtain from (48) and (42) that $v_i < \infty$ for every $i = 1, \ldots, m$. Since $\sum_{i=1}^{m} s_i^{\alpha} v_i^{1-\alpha} = 1$, we can apply Lemma D.2, (47), and (36) to obtain for $\alpha \in (-\infty, 1) \setminus \{0\}$ that

$$\liminf_{R\to\infty} e^{rR} D_{\mu}^{\alpha}(R) \geq C(r) \sum_{i=1}^{m} s_i t_i^{-r} l^r = C(r) \left(\int_M f^{a_1}\, d\lambda \right)^{a_2}.$$

Step 2.

Now let us assume that the support of μ is a compact interval $M \subset \mathbb{R}$ and that f is continuous on M. Again, let $\alpha \in (-\infty, 1) \setminus \{0\}$. Let $i(f) = \min\{f(x) : x \in M\}$ resp. $s(f) = \max\{f(x) : x \in M\}$. Clearly, $s(f) < \infty$. Let us assume that

$$i(f) > 0. \tag{49}$$

Let $l = \lambda(M)$. For $k \in \mathbb{N}$ partition M into intervals $\{A_i : i = 1, \ldots, k\}$ of common length $\lambda(A_i) = l/k$ Set

$$\mu_k = \sum_{i=1}^{k} \mu(A_i) U(A_i)$$

and

$$f_k = \frac{d\mu_k}{d\lambda} = \sum_{i=1}^{k} \frac{\mu(A_i)}{\lambda(A_i)} 1_{A_i}.$$

The continuity of f implies that f_k converges pointwise to f as $k \to \infty$. In view of (49) and due to

$$\begin{aligned} i(f) &= \min\{f(x) : x \in M\} \leq \min\{f_k(x) : x \in M\} \\ &\leq \max\{f_k(x) : x \in M\} \leq \max\{f(x) : x \in M\} = s(f) < \infty \end{aligned}$$

for every $k \in \mathbb{N}$, dominated convergence implies

$$\lim_{k\to\infty} \int_M f_k^{a_1}\, d\lambda = \int_M f^{a_1}\, d\lambda. \tag{50}$$

Moreover step 1 yields

$$\liminf_{R\to\infty} e^{rR} D_{\mu_k}^{\alpha}(R) \geq C(r) \left(\int_M f_k^{a_1}\, d\lambda \right)^{a_2}. \tag{51}$$

Now let $R \geq \max(1, R_0)$. Let $\delta > 0$ and q_R be a quantizer with $|D_\mu^\alpha(R) - D_\mu(q_R)| < \delta e^{-rR}$ and $H_\mu^\alpha(q_R) \leq R$. In addition, we assume w.l.o.g. (cf. Lemma 5.4, resp. Lemma 2.3) that $q_R \in \mathcal{K}_R(C)$ if $\alpha < 0$ and R is large enough, and $H_\mu^\alpha(q_R) = R$ if $\alpha > 0$. For $i = 1, \ldots, k$ let

$$0 < c_{i,k} = \min\{f(x) : x \in A_i\} \leq t_{i,k} = \max\{f(x) : x \in A_i\} < \infty$$

and

$$0 < c_k = \min\left\{\frac{c_{i,k}}{t_{i,k}} : i = 1, \ldots, k\right\}.$$

For every $a \in q_R(\mathbb{R})$ we have

$$c_k \mu_k(q_R^{-1}(a)) \leq \mu(q_R^{-1}(a)) \leq c_k^{-1} \mu_k(q_R^{-1}(a)).$$

and because f is uniformly continuous, in view of (49), we have

$$\lim_{k\to\infty} c_k = 1. \tag{52}$$

We obtain from the definitions of $H_\mu^\alpha(q_R)$ and c_k that

$$\min\left(c_k^{\frac{\alpha}{1-\alpha}}, c_k^{\frac{\alpha}{\alpha-1}}\right) \leq e^{H_\mu^\alpha(q_R) - H_{\mu_k}^\alpha(q_R)} \leq \max\left(c_k^{\frac{\alpha}{1-\alpha}}, c_k^{\frac{\alpha}{\alpha-1}}\right) =: v_k \tag{53}$$

where $v_k \to 1$ as $k \to \infty$. Again from the uniform continuity of f we deduce

$$\lim_{k\to\infty} \|f - f_k\|_\infty = 0 \tag{54}$$

with

$$\|f - f_k\|_\infty = \max\{|f(x) - f_k(x)| : x \in M\}.$$

In view of Proposition 5.1 there exists an $m_0 > 0$, such that for all $R \geq 1$

$$D_\mu^\alpha(R) \leq m_0 e^{-rR}. \tag{55}$$

By the choice of q_R we have

$$e^{rR} D_\mu^\alpha(R) \geq e^{rR} D_\mu(q_R) - \delta. \tag{56}$$

Thus (55) yields

$$\begin{aligned}
|D_\mu(q_R) - D_{\mu_k}(q_R)| &\leq \int_M |x - q_R(x)|^r |f(x) - f_k(x)| \, d\lambda(x) \\
&\leq \|f - f_k\|_\infty \frac{1}{i(f)} \int_M |x - q_R(x)|^r f(x) \, d\lambda(x) \\
&\leq \|f - f_k\|_\infty \frac{1}{i(f)} (D_\mu^\alpha(R) + \delta e^{-rR})
\end{aligned}$$

$$\leq \quad \|f-f_k\|_\infty \frac{m_0+\delta}{i(f)} e^{-rR}.$$

Hence, (56) gives

$$\begin{aligned}
&e^{rR} D_\mu^\alpha(R) \\
&\quad \geq \quad e^{rR}(D_{\mu_k}(q_R) - |D_\mu(q_R) - D_{\mu_k}(q_R)|) - \delta \\
&\quad \geq \quad e^{rR} D_{\mu_k}(q_R) - \|f-f_k\|_\infty \frac{m_0+\delta}{i(f)} - \delta \\
&\quad = \quad e^{r(R-H^\alpha_{\mu_k}(q_R))} e^{rH^\alpha_{\mu_k}(q_R)} D_{\mu_k}(q_R) - \|f-f_k\|_\infty \frac{m_0+\delta}{i(f)} - \delta. \\
&\quad \geq \quad e^{-r(|R-H^\alpha_\mu(q_R)|+|H^\alpha_\mu(q_R)-H^\alpha_{\mu_k}(q_R)|)} e^{rH^\alpha_{\mu_k}(q_R)} D_{\mu_k}(H^\alpha_{\mu_k}(q_R)) \\
&\qquad\quad - \|f-f_k\|_\infty \frac{m_0+\delta}{i(f)} - \delta. \qquad (57)
\end{aligned}$$

Due to the choice of q_R there exists a function $g : (0,\infty) \mapsto (0,\infty)$ with $e^{R-H^\alpha_\mu(q_R)} \leq g(R)$ and $g(R) \to 1$ as $R \to \infty$. Equation (53) implies

$$|H^\alpha_{\mu_k}(q_R) - R| \leq |H^\alpha_{\mu_k}(q_R) - H^\alpha_\mu(q_R)| + |H^\alpha_\mu(q_R) - R| \leq |\log(v_k)| + |\log(g(R))|. \qquad (58)$$

Clearly, inequality (58) yields $\lim_{R\to\infty} H^\alpha_{\mu_k}(q_R) = \infty$. Applying relations (58) and (51) to (57) we deduce

$$\liminf_{R\to\infty} e^{rR} D_\mu^\alpha(R) \geq e^{-r|\log(v_k)|} C(r) \left(\int_M f_k^{a_1}\, d\lambda \right)^{a_2} - \|f-f_k\|_\infty \frac{m_0+\delta}{i(f)} - \delta.$$

By letting $k \to \infty$ and noting that $\delta > 0$ is arbitrary we obtain from (50), (52) and (54) that

$$\liminf_{R\to\infty} e^{rR} D_\mu^\alpha(R) \geq C(r) \left(\int_M f^{a_1}\, d\lambda \right)^{a_2}. \qquad (59)$$

Next we show a matching upper bound for $\alpha \in (-\infty, 0)$. The assumptions on f allow us to use Corollary 4.11 showing the existence of a sequence of companding quantizers $(q_N) = \left(Q_{f^*,N}\right)$ such that

$$\lim_{N\to\infty} e^{rH^\alpha_\mu(q_N)} D_\mu(q_N) \leq C(r) \left(\int_M f^{a_1}\, d\lambda \right)^{a_2}. \qquad (60)$$

Let $R_N = H^\alpha_\mu(q_N)$ and note that Proposition 4.8 implies

$$\lim_{N\to\infty} R_N = \infty, \quad \lim_{N\to\infty} (R_N - R_{N-1}) = 0. \qquad (61)$$

Let $R > 0$ be arbitrary and let $n = \max\{N : R_N \leq R\}$. Then $R_n \leq R < R_{n+1}$ and since $D^\alpha_\mu(R)$ is a

nonincreasing function of R

$$e^{rR}D_\mu^\alpha(R) \leq e^{rR_{n+1}}D_\mu^\alpha(R_n) \leq e^{r(R_{n+1}-R_n)}e^{rR_n}D_\mu(q_N).$$

This, (60), and (61) yield

$$\limsup_{R\to\infty} e^{rR}D_\mu^\alpha(R) \leq C(r)\left(\int_M f^{a_1}\,d\lambda\right)^{a_2}.$$

Together with (59) this completes the proof for the case $\alpha \in (-\infty, 0)$.
Step 3.

Now let μ be arbitrary, but satisfying all assumptions of the theorem. Let $\alpha \in (0,1)$. For $k, l \in \mathbb{N}$ let

$$I_1 = (-\infty, -k), \quad I_2 = [-k,k] \cap f^{-1}([1/l, l]), \quad I_3 = (k, \infty)$$

and

$$I_4 = \mathbb{R} \setminus (I_1 \cup I_2 \cup I_3).$$

Because f is bounded and weakly unimodal we can pick $k_0 \in \mathbb{N}$ such that $\mu(I_2) > 0$, $f^{-1}([1/l,l]) = f^{-1}([1/l,\infty))$, $1/l < l_0$ (see Definition 3.2), and

$$\mu(I_2) = \max\{\mu(I_i) : i \in \{1,2,3,4\}\}$$

for every $k \geq k_0$ and $l \geq k_0$. Note that I_2 is a compact interval. Now let $\min(k,l) \geq k_0$. Let us first assume that $\mu(I_i) > 0$ for every $i = 1,2,3,4$. Consider the decomposition $\mu = \sum_{i=1}^4 \mu(I_i)\mu(\cdot|I_i)$. Lemma 6.3 yields

$$\liminf_{R\to\infty} e^{rR}D_\mu^\alpha(R) \geq \mu(I_2)^{\frac{1-\alpha+\alpha r}{1-\alpha}} \liminf_{R\to\infty} e^{rR}D^\alpha_{\mu(\cdot|I_2)}(R). \tag{62}$$

By construction, $i(\mu(I_2)^{-1}f1_{I_2}) > 0$, $s(\mu(I_2)^{-1}f1_{I_2}) < \infty$, and $\mu(I_2)^{-1}f1_{I_2}$ is supported by a compact interval. Thus we can apply the results of step 2. Together with the definition of a_1 and a_2 we deduce from (59) and (62) that

$$\begin{aligned}
\liminf_{R\to\infty} e^{rR}D_\mu^\alpha(R) &\geq \mu(I_2)^{\frac{1-\alpha+\alpha r}{1-\alpha}}C(r)\left(\int_{I_2}\left(\mu\big(I_2\big)^{-1}f\right)^{a_1}d\lambda\right)^{a_2} \\
&\geq \mu(I_2)^{\frac{1-\alpha+\alpha r}{1-\alpha}-a_1a_2}C(r)\left(\int_{I_2} f^{a_1}\,d\lambda\right)^{a_2} \\
&= C(r)\left(\int_{I_2} f^{a_1}\,d\lambda\right)^{a_2}. \qquad (63)
\end{aligned}$$

Due to $a_1 > 0$ and by monotone convergence we obtain

$$\lim_{k\to\infty}\lim_{l\to\infty}\int_{I_2} f^{a_1}\,d\lambda = \int f^{a_1}\,d\lambda. \tag{64}$$

Thus we get from (63) that

$$\liminf_{R\to\infty} e^{rR} D_\mu^\alpha(R) \geq C(r)\left(\int f^{a_1}\,d\lambda\right)^{a_2}. \tag{65}$$

The case $\min\{\mu(I_1), \mu(I_3), \mu(I_4)\} = 0$ can be treated similarly.

To show the matching upper bound, note that since $1/l \leq f \leq l$ on I_2, we can directly apply Corollary 4.11 to $\mu_2 := \mu(\cdot|I_2)$ and its density $f_2 := \mu(I_2)^{-1} f 1_{I_2}$ to show the existence of a sequence of companding quantizers $(q_N) = \left(Q_{f_2^*,N}\right)$ such that

$$\lim_{N\to\infty} e^{rH^\alpha_{\mu_2}(q_N)} D_{\mu_2}(q_N) \leq C(r)\left(\int f_2^{a_1}\,d\lambda\right)^{a_2}.$$

Thus by the same argument as in the previous step

$$\begin{aligned}\limsup_{R\to\infty} e^{rR} D^\alpha_{\mu_2}(R) &\leq C(r)\left(\int f_2^{a_1}\,d\lambda\right)^{a_2} \\ &= C(r)\mu(I_2)^{-a_1a_2}\left(\int_{I_2} f^{a_1} d\lambda\right)^{a_2}.\end{aligned} \tag{66}$$

Again from Lemma 6.3 we obtain for $\alpha \in (0,1)$ the upper bound

$$\begin{aligned}\limsup_{R\to\infty} e^{rR} D^\alpha_\mu(R) &\leq C(r)\mu(I_2)^{1-a_1a_2} t_2^{-r}\left(\int_{I_2} f^{a_1} d\lambda\right)^{a_2} \\ &\quad + 3\max_{i=1,3,4}\mu(I_i)t_i^{-r}\limsup_{R\to\infty} e^{rR} D^\alpha_{\mu(\cdot|I_i)}(R).\end{aligned}$$

Using Proposition 5.1 we get a $K > 0$ independent of k, l such that

$$\limsup_{R\to\infty} e^{rR} D^\alpha_{\mu(\cdot|I_i)}(R) \leq K\mu(I_i)^{-r/(r+\delta)}.$$

for $i \in \{1,3,4\}$. Letting l, k tend to infinity we obtain by the definition of t_i that $\lim_{l,k\to\infty} t_2^{-r} = 1$, resp. $\lim_{l,k\to\infty} t_i^{-r} = 0$, $i = 1,3,4$. Using (64) and (66) we get

$$\limsup_{R\to\infty} e^{rR} D^\alpha_\mu(R) \leq C(r)\left(\int f^{a_1} d\lambda\right)^{a_2}$$

which, together with the lower bound (65) completes the proof for the case $\alpha \in (0,1)$.
Step 4.

Let $\alpha = -\infty$ and $\beta \in (-\infty, 0)$. Fix $a_1 = a_1(\beta)$ and $a_2 = a_2(\beta)$. From Lemma 2.3 we deduce

$$\liminf_{R\to\infty} e^{rR} D_\mu^{-\infty}(R) \geq \lim_{R\to\infty} e^{rR} D_\mu^{\beta}(R) = C(r) \left(\int_{\mathrm{supp}(\mu)} f^{a_1}\, d\lambda \right)^{a_2}.$$

Since the integral on the right hand side converges to $\int f^{1-r}\, d\lambda$ as $\beta \to -\infty$, we obtain

$$\liminf_{R\to\infty} e^{rR} D_\mu^{-\infty}(R) \geq C(r) \int_{\mathrm{supp}(\mu)} f^{1-r}\, d\lambda.$$

The proof is finished by noting that Corollary 4.11 and an argument identical to the one used in step 2 provide a matching upper bound. □

8 Concluding remarks

We have determined the sharp distortion asymptotics for optimal scalar quantization with Rényi entropy constraint for values $\alpha \in [-\infty, 0) \cup (0, 1)$ of the order parameter. Our results, together with the classical $\alpha = 0$ and $\alpha = 1$ cases, and the recent result [18] for $\alpha \in [r+1, \infty]$, leave only open the case $\alpha \in (1, 1+r)$ for which non-matching upper and lower bounds are known to date (cf. [18]). We note that the upper bound provided by optimal companding in Corollary 4.11 also holds for $\alpha \in (1, 1+r)$. Based on this, we conjecture that our main result is also valid for this remaining range of the α parameter.

Apart from the question of high-rate asymptotics, it remains open if optimal quantizers exist for all $\alpha \in [-\infty, 1+r]$. The non-existence of optimal quantizers in case of $\alpha > 1+r$ has already been shown in [19]. Looking at our main result, it is obvious that the integrals on the right hand sides of (11) and (12) are not finite in general if μ has unbounded support. It needs further research to determine the exact high-rate error asymptotics for certain classes of source distributions with unbounded support and $\alpha < 0$. Of special interest is the question whether companding quantizers with point density f^* are still asymptotically optimal for source densities with unbounded support. The definition of f^* needs the integrability of f^{1/a_2} in order to guarantee a finite number of quantization points for the (asymptotically optimal) companding quantizer. Nevertheless, the right hand side of (11) is defined only when f^{a_1} is integrable. It remains an open problem if (11) still holds for some $\alpha \in (-\infty, 1+r)\setminus\{1\}$ and distributions where f^{a_1} is integrable but f^{1/a_2} is not. Such an example, if it exists, would show that the companding approach is not always applicable to generate asymptotically optimal quantizers, but the known asymptotics (11) are still in force. Another interesting open question is whether the non-integrability of f^{1/a_2} always implies the non-existence of optimal quantizers with a finite codebook.

A careful reading of the proofs shows that many arguments can be straightforwardly generalized to the d-dimensional case and rth power distortion based on some norm on $\mathbb{R}^d$. For $\alpha \in [-\infty, 1)$ and under appropriate conditions we conjecture that

$$\limsup_{R\to\infty} e^{\frac{r}{d}R} D_\mu^{\alpha}(R) = C(r, d) \left(\int f^{a_1}\, d\lambda^d \right)^{a_2}$$

where λ^d is the d-dimensional Lebesgue measure,

$$a_1 = \frac{1-\alpha+\alpha\frac{r}{d}}{1-\alpha+\frac{r}{d}}, \quad a_2 = \frac{1-\alpha+\frac{r}{d}}{1-\alpha}$$

and $C(r,d)$ is a positive constant that depends only on r, d, and the underlying norm.

However, some important steps in our proofs are definitely restricted to the scalar case, e.g., equation (23) in Lemma 5.3, which yields (40). One of the key problems concerns the first step of the proof of Theorem 3.4. In higher dimensions one has to control the contribution to distortion and entropy of cells straddling the common boundary of at least two touching cubes in the support of μ. The “firewall” construction used in case of $\alpha = 0$ (see [9, p.87]) does not seem to work in the general case. For $\alpha \neq 0$ it seems to be very hard to control the entropy of the quantizer when adding or changing codecells and codepoints in a certain region. In order to progress in this direction, one would certainly need more refined knowledge about the codecell geometry of (asymptotically) optimal quantizers. Even in the case $\alpha = 0$ little is known on this subject (results in [29] highlight the difficulty of the problem). As already mentioned in the introduction, the methods used for the case $\alpha = 1$ are also not applicable to the general case because they rely on the special functional form of the Shannon entropy. It appears that generalization to higher dimensions would necessitate the development of isodiametric inequalities for the (bounded) codecells of asymptotically optimal quantizers.

Appendix A

Proof of Lemma 2.3. To show (4), let $q \in \mathcal{Q}$ be such that $H^{\alpha}_{\mu}(q) \leq R$ and assume $\beta \leq \alpha$. It is easy to check that $\frac{d}{d\gamma} H^{\gamma}_{\mu}(q) \leq 0$ on $(-\infty, 0) \cup (0,1) \cup (1,\infty)$, and thus the mapping $\gamma \mapsto H^{\gamma}_{\mu}(q)$ is non-increasing on these intervals. In view of the continuity of $H^{\alpha}_{\mu}(q)$ at $\alpha \in \{0,1\}$ (see (Remark 2.2(a)) we deduce that $H^{\alpha}_{\mu}(q) \leq H^{\beta}_{\mu}(q)$. Now the assertion follows from Definition (3).

Equation (5) of the second statement follows directly from the more general results Theorem 3.2 and Proposition 4.2 in [20]. For (6), we refer to [19, Proposition 2.1.(i)]. □

Proof of Proposition 4.2. We proceed in several steps.

1. Since $\hat{G}$ is increasing, it has a derivative $\hat{G}'$ a.e. (by convention we set $\hat{G}'(x) = 0$ if $\hat{G}$ is not differentiable at x). Also, note that G and $\hat{G}$ are strictly increasing on I, resp. on $[0,1]$, and $\hat{G}$ is Lipschitz with constant $(\operatorname{ess\,inf}_I g)^{-1}$ and thus absolutely continuous. Since $\hat{G}'(x) = 1/g(\hat{G}(x))$ a.e. on $(0,1)$, we obtain

$$\int \frac{f}{g^r}\, d\lambda = \int (\hat{G}'(G(x)))^r \, d\mu(x).$$

2. Next we prove

$$\lim_{N\to\infty} \int (m_N(G(x)))^r \, d\mu(x) = \int (\hat{G}'(G(x)))^r \, d\mu(x), \tag{67}$$

where m_N is the piecewise constant function defined by $m_N(x) = N \int_{[(i-1)/N, i/N)} \hat{G}' \, d\lambda$, if $x \in [(i-1)/N, i/N)$, $i = 1, \ldots, N$ and $m_N(x) = 0$ otherwise.

Lebesgue's differentiation theorem (see, e.g., [7, Thm. 6.2.3]) implies that $m_N(x) \to \hat{G}'(x)$ as $N \to \infty$ a.e. on $(0, 1)$. Also, from $\hat{G}'(x) = 1/g(\hat{G}(x))$ we deduce

$$\begin{aligned} m_N(\cdot) &\leq \max\left\{ N \int_{[(i-1)/N, i/N)} \frac{1}{g(\hat{G}(x))} \, d\lambda(x) : 1 \leq i \leq N \right\} \\ &\leq (\operatorname{ess\,inf}_I g)^{-1}. \end{aligned}$$

Thus, the dominated convergence theorem yields (67).

3. Let $\bar{Q}_{g,N}$ be the quantizer with the same codecells as $Q_{g,N}$ but with the midpoints of the codecells as quantization points:

$$\bar{Q}_{g,N}(x) = \frac{1}{2}(\hat{G}(i/N) + \hat{G}((i-1)/N)) \quad \text{if } x \in I_{i,N}, \; i = 1, \ldots, N.$$

We will show that

$$\lim_{N \to \infty} N^r \int |x - \bar{Q}_{g,N}(x)|^r f_N(x) \, d\lambda(x) = C(r) \int \frac{f}{g^r} \, d\lambda,$$

where f_N is the piecewise constant density defined by $f_N(x) = \frac{1}{\lambda(I_{i,N})} \int_{I_{i,N}} f \, d\lambda$ if $x \in I_{i,N}$, $i = 1, \ldots, N$ and $f_N(x) = 0$ otherwise.

A simple calculation shows

$$\begin{aligned} & N^r \int |x - \bar{Q}_{g,N}(x)|^r f_N(x) \, d\lambda(x) \\ &\quad = N^r \sum_{i=1}^{N} C(r) f_N\left(\hat{G}\left(\frac{i}{N}\right)\right) (\lambda(I_{i,N}))^{r+1} \\ &\quad = C(r) N^r \sum_{i=1}^{N} \int_{I_{i,N}} f(x) (N^{-1} m_N(G(x)))^r \, d\lambda(x) \\ &\quad = C(r) \int m_N(G(x))^r f(x) \, d\lambda(x). \end{aligned}$$

Now the assertion follows from steps 1 and 2.

4. Next we show that

$$\lim_{N \to \infty} N^r D_\mu(\bar{Q}_{g,N}) = C(r) \int \frac{f}{g^r} \, d\lambda.$$

For any $i \in \{1, \ldots, N\}$ and $x \in I_{i,N}$ we have

$$N^r |x - \bar{Q}_{g,N}(x)|^r \leq N^r \big(\lambda(I_{i,N})\big)^r = N^r \left(\int_{((i-1)/N, i/N)} \hat{G}' \, d\lambda \right)^r = m_N(G(x))^r.$$

Therefore

$$\begin{aligned}
&\left| N^r \int |x - \bar{Q}_{g,N}(x)|^r f(x)\, d\lambda(x) - N^r \int |x - \bar{Q}_{g,N}(x)|^r f_N(x)\, d\lambda(x) \right| \\
&\quad \leq \int m_N(G(x))^r |f(x) - f_N(x)|\, d\lambda(x) \\
&\quad \leq (\operatorname{ess\,inf}_I g)^{-1} \int |f(x) - f_N(x)|\, d\lambda(x).
\end{aligned}$$

By Lebesgue's differentiation theorem we have $f_N \to f$ a.e., and now Scheffé's theorem [4, Thm. 16.11] implies

$$\lim_{N\to\infty} \int |f - f_N|\, d\lambda = 0.$$

Hence,

$$\lim_{N\to\infty} N^r D_\mu(\bar{Q}_{g,N}) = \lim_{N\to\infty} N^r \int |x - \bar{Q}_{g,N}(x)|^r f_N(x)\, d\lambda(x),$$

where the right hand side is equal to $C(r) \int \frac{f}{g^r}\, d\lambda$ from step 3.

5. In view of step 4, to prove relation (16) it suffices to show that

$$\lim_{N\to\infty} N^r D_\mu(\bar{Q}_{g,N}) = \lim_{N\to\infty} N^r D_\mu(Q_{g,N}). \tag{68}$$

Applying the mean value theorem of differentiation (if $r > 1$) or by the triangle inequality (if $r = 1$), we have for each $i \in \{1, \ldots, N\}$ and $x \in I_{i,N}$,

$$\begin{aligned}
&N^r \big| |x - \bar{Q}_{g,N}(x)|^r - |x - Q_{g,N}(x)|^r \big| \\
&\quad \leq N^r |\bar{Q}_{g,N}(x) - Q_{g,N}(x)| r (\lambda(I_{i,N}))^{r-1}.
\end{aligned} \tag{69}$$

Further, note that the definitions of m_N, $Q_{g,N}$, $\bar{Q}_{g,N}$ also yield

$$\begin{aligned}
&N^r \big| |x - \bar{Q}_{g,N}(x)|^r - |x - Q_{g,N}(x)|^r \big| \\
&\quad \leq r N^r (\lambda(I_{i,N}))^r = r \cdot m_N(G(x))^r.
\end{aligned} \tag{70}$$

Let $I_{i,N}^1$ and $I_{i,N}^2$ denote the partition of $I_{i,N}$ into two intervals of equal length $\lambda(I_{i,N})/2$. Let $j = j(x) \in \{1, 2\}$ be such that $x \in I_{i,N}^j$. Letting $a = \inf(I_{i,N})$ and $b = \sup(I_{i,N})$ we obtain by the absolute continuity of $\hat{G}$

$$\hat{G}((a+b)/2) = \hat{G}(a) + \int_a^{(a+b)/2} \hat{G}'\, d\lambda$$

and

$$\frac{\hat{G}(a) + \hat{G}(b)}{2} = \hat{G}(a) + \int_a^{(a+b)/2} \frac{\hat{G}(b) - \hat{G}(a)}{b - a}\, d\lambda.$$

Thus we get

$$
\begin{aligned}
|\bar{Q}_{g,N}(x) - Q_{g,N}(x)| &= \left|\hat{G}((a+b)/2) - \frac{\hat{G}(a)+\hat{G}(b)}{2}\right| \\
&\leq \lambda(I^j_{i,N})|L(x,N)| \qquad (71)
\end{aligned}
$$

where

$$
\begin{aligned}
L(x,N) &= \frac{1}{\lambda(I^j_{i,N})}\int_{I^j_{i,N}} \hat{G}'\,d\lambda - \frac{\hat{G}(b)-\hat{G}(a)}{b-a} \\
&= \frac{1}{\lambda(I^j_{i,N})}\int_{I^j_{i,N}} \hat{G}'\,d\lambda - \frac{1}{\lambda(I_{i,N})}\int_{I_{i,N}} \hat{G}'\,d\lambda
\end{aligned}
$$

if $x \in I^j_{i,N}$, $i = 1, \dots, N$. In view of (69) and (71) we deduce

$$
\begin{aligned}
N^r\big||x-\bar{Q}_{g,N}(x)|^r - |x - Q_{g,N}(x)|^r\big| &\leq rN^r\lambda(I^j_{i,N})|L(x,N)|(\lambda(I_{i,N}))^{r-1} \\
&\leq r \cdot m_N(G(x))^r|L(x,N)|.
\end{aligned}
$$

Lebesgue's differentiation theorem yields $\lim_{N\to\infty} L(x,N) = 0$ a.e. Hence

$$
\lim_{N\to\infty} N^r\big||x-\bar{Q}_{g,N}(x)|^r - |x - Q_{g,N}(x)|^r\big| = 0 \quad \text{a.e.} \qquad (72)
$$

Due to the relations (72), (70) and together with step 2, we can apply the generalized dominated convergence theorem [28, Chapter 11.4] to obtain (68). □

Lemma A.1. *Let μ be a probability distribution which is absolutely continuous with respect to λ and let f denote its density. Let G be a compressor for μ with point density g. If $\{g = 0\} \subset \{f = 0\}$, then $\mu \circ G^{-1}$ is absolutely continuous with respect to λ. Also*

$$
\hat{G}'(y) = \frac{1}{g(\hat{G}(y))}1_{\{g>0\}}(\hat{G}(y)) \quad \textit{a.e. on } (0,1) \qquad (73)
$$

and $\mu \circ G^{-1}$ has the density

$$
f_G(y) = f(\hat{G}(y))\hat{G}'(y) = \frac{f(\hat{G}(y))}{g(\hat{G}(y))}1_{\{g>0\}}(\hat{G}(y)) \quad \textit{a.e. on } (0,1). \qquad (74)
$$

Proof. In order to prove that $\mu \circ G^{-1}$ is absolutely continuous let us make the key observation that, although G is in general not invertible, we have

$$
\hat{G}(G(x)) = x \quad \mu\text{-a.e. } x \in \mathbb{R} \qquad (75)
$$

Indeed, by the definition of G and due to $\{g = 0\} \subset \{f = 0\}$ there exists a measurable set $A_G \subset \mathbb{R}$

such that G is differentiable on A_G, $\mu(A_G) = 1$, and

$$G'(x) = g(x) \in (0, \infty) \text{ for every } x \in A_G.$$

Hence G is locally invertible at $x \in A_G$, so $\hat{G}(G(x)) = x$ which proves (75). Moreover,

$$\hat{G}'(G(x)) = 1/g(x) \text{ for every } x \in A_G \tag{76}$$

which proves (73).

$\hat{G}$ is strictly increasing (and thus one-to-one) and maps $(0,1)$ onto $\hat{G}((0,1))$. Thus, together with (75) we obtain for every Borel measurable $B \subset \mathbb{R}$ that

$$\mu \circ G^{-1}(B) = \mu(\{x : \hat{G}(G(x)) \in \hat{G}(B)\}) = \mu(\hat{G}(B)). \tag{77}$$

If $U([0,1])$ denotes the uniform distribution on $[0,1]$ we obtain again from (75) that

$$U([0,1]) \circ \hat{G}^{-1}((-\infty, x]) = G(x) \quad \text{for a.e. } x \in \mathbb{R}.$$

Thus, $U([0,1]) \circ \hat{G}^{-1} = g\lambda$. Now let $B \subset (0,1)$ be Borel measurable and $\lambda(B) = 0$. This implies $U([0,1]) \circ \hat{G}^{-1}(\hat{G}(B)) = 0$. Because μ is absolutely continuous with respect to $g\lambda$ we obtain $\mu(\hat{G}(B)) = 0$. Hence, (77) implies $\mu \circ G^{-1}(B) = 0$ showing that $\mu \circ G^{-1}(B)$ is absolutely continuous with respect to λ. In order to prove (74) let $[a,b] \subset (0,1)$. In view of (77) and from the definition of $\hat{G}$ we obtain

$$\mu(G^{-1}([a,b])) = \int_{\hat{G}(a)}^{\hat{G}(b)} f \, d\lambda. \tag{78}$$

From (76) we deduce

$$\hat{G}'(y) = \frac{1}{g(\hat{G}(y))} 1_{\{g>0\}}(\hat{G}(y)) \quad \text{a.e. on } (0,1).$$

Because $\mu \circ G^{-1}$ is absolutely continuous with respect to λ its cumulative distribution function is absolutely continuous and, therefore, differentiable a.e. Applying the chain rule for the Lebesgue integral (see [30, Corollary 4]) we obtain

$$\int_{\hat{G}(a)}^{\hat{G}(b)} f(x) \, d\lambda(x) = \int_a^b f(\hat{G}(y)) \hat{G}'(y) \, d\lambda(y). \tag{79}$$

Now, (78) and (79) prove the first equation in (74). The second equality in (74) follows from (73). $\square$

Proof of Lemma 4.6.
1. $\alpha \in (0, \infty)$. For this range of α the result goes back to Rényi [27, 11§] who stated it with somewhat less generality. Csiszár [8, Thm. 2] gives a more general form of the result that implies our statement.

2. $\alpha = -\infty$.

Clearly,

$$\liminf_{\Delta\to 0} \frac{\inf\{\mu(\hat{q}_{\Delta,M}^{-1}(a)) : a \in \hat{q}_{\Delta,M}(\mathbb{R})\}}{\Delta} \geq \operatorname{ess\,inf}_M f. \tag{80}$$

Now let $\varepsilon > 0$ and define $N_\varepsilon = \{x : f(x) < \operatorname{ess\,inf}_M f + \varepsilon\} \cap M$. Hence, $\lambda(N_\varepsilon) > 0$. By Lebesgue's differentiation theorem we can find an $x \in N_\varepsilon$ such that μ is differentiable at x with $f(x) = \frac{d\mu}{d\lambda}(x)$. Moreover a $\Delta_0(\varepsilon) > 0$ exists, such that for every $\Delta \leq \Delta_0$ a $b \in \hat{q}_{\Delta,M}(\mathbb{R})$ can be found with $x \in \hat{q}_{\Delta,M}^{-1}(b)$ and

$$\frac{\mu(\hat{q}_{\Delta,M}^{-1}(b))}{\Delta} \leq \operatorname{ess\,inf}_M f + 2\varepsilon.$$

Because ε is arbitrary we obtain

$$\limsup_{\Delta\to 0} \frac{\inf\{\mu(\hat{q}_{\Delta,M}^{-1}(a)) : a \in \hat{q}_{\Delta,M}(\mathbb{R})\}}{\Delta} \leq \operatorname{ess\,inf}_M f. \tag{81}$$

In view of Definition 4.4 and the definition of $H_\mu^{-\infty}(\cdot)$, the combination of (80) and (81) yields the assertion.

3. $\alpha \in (-\infty, 0]$. Here we adapt Rényi's original proof to our case. With the convention $0^0 := 0$ and in view of Definition 4.4 resp. Remark 2.2 it suffices to show that

$$\int_{\operatorname{supp}(\mu)} f^\alpha \, d\lambda = \lim_{\Delta\to 0} \sum_{a\in\hat{q}_\Delta(\mathbb{R})} \Delta^{1-\alpha} \left(\int_{\hat{q}_\Delta^{-1}(a)} f \, d\lambda \right)^\alpha. \tag{82}$$

For $\Delta > 0$ and $x \in \mathbb{R}$ we define

$$g_{1,\Delta}(x) = 1_{\operatorname{supp}(\mu)}(x) \sum_{a\in\hat{q}_\Delta(\mathbb{R})} 1_{\hat{q}_\Delta^{-1}(a)}(x) \frac{1}{\Delta} \int_{\hat{q}_\Delta^{-1}(a)} f^\alpha \, d\lambda$$

and

$$\begin{aligned} g_{2,\Delta}(x) &= 1_{\operatorname{supp}(\mu)}(x) \sum_{a\in\hat{q}_\Delta(\mathbb{R})} 1_{\hat{q}_\Delta^{-1}(a)}(x) \Delta^{-\alpha} \left(\int_{\hat{q}_\Delta^{-1}(a)} f \, d\lambda \right)^\alpha \\ &= 1_{\operatorname{supp}(\mu)}(x) \left(\sum_{a\in\hat{q}_\Delta(\mathbb{R})} 1_{\hat{q}_\Delta^{-1}(a)}(x) \frac{1}{\Delta} \int_{\hat{q}_\Delta^{-1}(a)} f \, d\lambda \right)^\alpha. \end{aligned}$$

Applying Lebesgue's differentiation theorem we obtain $g_{1,\Delta} \to f^\alpha$ and $g_{2,\Delta} \to f^\alpha$ a.e. as $\Delta \to 0$. Now note that since $\alpha \leq 0$, the function $x \mapsto x^\alpha$ is convex on $(0, \infty)$, so by Jensen's inequality

$$\left(\frac{1}{\Delta} \int_{\hat{q}_\Delta^{-1}(a)} f \, d\lambda \right)^\alpha \leq \frac{1}{\Delta} \left(\int_{\hat{q}_\Delta^{-1}(a)} f^\alpha \, d\lambda \right)$$

for all $a \in q^{-1}(\mathbb{R})$, implying $g_{2,\Delta}(x) \leq g_{1,\Delta}(x)$ for all x. Since $g_{2,\Delta} \geq 0$ and $\int g_{1,\Delta}\, d\lambda = \int_{\mathrm{supp}(\mu)} f^\alpha\, d\lambda \in (0,\infty)$, we can apply the generalized dominated convergence theorem [28, Chapter 11.4] to obtain

$$\lim_{\Delta \to 0} \int g_{1,\Delta}\, d\lambda = \lim_{\Delta \to 0} \int g_{2,\Delta}\, d\lambda,$$

which is equivalent to (82) and, therefore, finishes the proof. □

Appendix B

Proof of Proposition 5.1.

(i) Recall that $\lfloor x \rfloor$ denotes the largest integer less than or equal to $x \in \mathbb{R}$. In view of Lemma 2.3 we have $D^\alpha_\mu(R) \leq D^0_\mu(R)$. Consequently, we deduce from [23, Lemma 1] the existence of a constant $\kappa > 0$ (that depends only on r and δ) such that

$$D^\alpha_\mu(R) \leq D^0_\mu(R) \leq D^0_\mu(\log(\lfloor e^R \rfloor)) \leq (\lfloor e^R \rfloor)^{-r} \kappa^r \left(\int |x|^{r+\delta}\, d\mu(x) \right)^{r/(r+\delta)}.$$

Due to $R \geq 1$ we obtain

$$(\lfloor e^R \rfloor)^{-r} \leq e^{-rR} \left(\frac{e^R}{e^R - 1} \right)^r \leq e^{-rR} \left(\frac{e}{e-1} \right)^r,$$

which yields the assertion with $C_0 = \kappa^r \left(\frac{e}{e-1} \right)^r$.

(ii) In view of Lemma 2.3 it is enough to prove relation (21) for $\alpha = -\infty$. Let $I = \mathrm{supp}(\mu)$, $R \geq 0$ and $q_R \in \mathcal{Q}$ with $H^{-\infty}_\mu(q_R) \leq R$. According to Lemma 2.3 let us assume w.l.o.g. that all codecells of q_R with positive $\mu-$mass are intervals. By subdivision of codecells with $\mu-$mass greater than or equal to $2e^{-R}$ we can assume w.l.o.g. that

$$e^{-R} \leq \mu(q_R^{-1}(a)) < 2e^{-R} \tag{83}$$

for every $a \in q_R(\mathbb{R})$ with $\mu(q_R^{-1}(a)) > 0$, where the first inequality holds since $H^{-\infty}_\mu(q_R) \leq R$. Moreover, for every such a we obtain

$$\mathrm{diam}(q_R^{-1}(a) \cap I) \leq \frac{\mu(q_R^{-1}(a))}{i(f)} \tag{84}$$

and we can assume w.l.o.g. that $a \in q_R^{-1}(a) \cap I$ if $q_R^{-1}(a)) \cap I \neq \emptyset$ (otherwise the distortion can be decreased by redefining a). Then we have

$$D_\mu(q_R) \quad = \quad \sum_{a \in q_R(\mathbb{R})} \int_{q_R^{-1}(a)} |x-a|^r f(x)\, d\lambda(x)$$

$$\leq \sum_{a\in q_R(\mathbb{R})} \int_{q_R^{-1}(a)} (\operatorname{diam}(q_R^{-1}(a)\cap I))^r f(x)\,d\lambda(x).$$

In view of (83) and (84) we get

$$\begin{aligned} D_\mu(q_R) &\leq \sum_{a\in q_R(\mathbb{R})} \int_{q_R^{-1}(a)} \left(\frac{\mu(q_R^{-1}(a))}{i(f)}\right)^r f(x)\,d\lambda(x) \\ &= \frac{1}{i(f)^r} \sum_{a\in q_R(\mathbb{R})} (\mu(q_R^{-1}(a)))^{r+1} \\ &< \frac{1}{i(f)^r} \sum_{a\in q_R(\mathbb{R})} \mu(q_R^{-1}(a))(2e^{-R})^r = \frac{2^r}{i(f)^r} e^{-rR}, \end{aligned}$$

which yields (21) by taking the infimum over all $q_R \in \mathcal{Q}$ with $H_\mu^{-\infty}(q_R) \leq R$. ☐

Proof of Lemma 5.3. Let $\varepsilon > 0$. Choose $c, t \in (0,\infty)$, such that

$$1 - \mu(A_{c,t}) < \frac{\varepsilon}{2} \tag{85}$$

with $A_{c,t} = \{x : f(x) \geq c\} \cap \{x : f(x) \leq t\} \cap [-t,t]$. Let

$$\kappa = \frac{c}{(1+r)2^r}$$

and use Corollary 5.2 to choose $R_0 > 0$ such that

$$t\left(\frac{2D_\mu^\alpha(R_0)}{\kappa}\right)^{\frac{1}{1+r}} < \frac{\varepsilon}{2}.$$

Now let $R \geq R_0$, $\delta = D_\mu^\alpha(R) > 0$, and choose $q \in \mathcal{Q}$ with $H_\mu^\alpha(q) \leq R$ and $|D_\mu(q) - D_\mu^\alpha(R)| < \delta$. We have

$$\begin{aligned} D_\mu(q) &= \sum_{a\in q(\mathbb{R})} \int_{q^{-1}(a)} |x-a|^r f(x)\,d\lambda(x) \\ &\geq \sum_{a\in q(\mathbb{R})} c \int_{q^{-1}(a)\cap A_{c,t}} |x-a|^r\,d\lambda(x). \end{aligned} \tag{86}$$

Let $B(x,l) = [x-l, x+l]$ for any $l > 0$ and $x \in \mathbb{R}$. For every $a \in q(\mathbb{R})$ define

$$s_a = \lambda(q^{-1}(a)\cap A_{c,t})/2. \tag{87}$$

Since $A_{c,t}$ is bounded, we have $s_a \in [0, \infty)$. Moreover, it is easy to show that

$$\int_{q^{-1}(a)\cap A_{c,t}} |x-a|^r \, d\lambda(x) \geq \int_{B(a,s_a)} |x-a|^r \, d\lambda(x) \tag{88}$$

(see, e.g., [9, Lemma 2.8]). Using (87) we compute

$$\int_{B(a,s_a)} |x-a|^r \, d\lambda(x) = \frac{2s_a^{r+1}}{1+r} = \frac{s_a^r}{1+r}\lambda(q^{-1}(a) \cap A_{c,t}). \tag{89}$$

Combining (88) and (89) with (86) we obtain

$$\begin{aligned} D_\mu(q) &\geq \frac{c}{1+r} \sum_{a\in q(\mathbb{R})} \lambda(q^{-1}(a) \cap A_{c,t}) s_a^r \\ &= \frac{c}{(1+r)2^r} \sum_{a\in q(\mathbb{R})} \lambda(q^{-1}(a) \cap A_{c,t})(2s_a)^r. \end{aligned}$$

Using (87) we get

$$\begin{aligned} D_\mu(q) &\geq \frac{c}{(1+r)2^r} \sum_{a\in q(\mathbb{R})} \lambda(q^{-1}(a) \cap A_{c,t})^{1+r} \\ &\geq \kappa \cdot \sup_{a\in q(\mathbb{R})} \lambda(q^{-1}(a) \cap A_{c,t})^{1+r}. \end{aligned}$$

On the other hand the choice of δ and the monotonicity of $D_\mu^\alpha(\cdot)$ yields

$$D_\mu(q) \leq 2D_\mu^\alpha(R) \leq 2D_\mu^\alpha(R_0)$$

Thus we deduce

$$\kappa \cdot \sup_{a\in q(\mathbb{R})} \lambda(q^{-1}(a) \cap A_{c,t})^{1+r} \leq 2D_\mu^\alpha(R_0).$$

Also, since f is upper bounded by t on $A_{c,t}$,

$$\begin{aligned} \max_{a\in q(\mathbb{R})} \mu(q^{-1}(a) \cap A_{c,t}) &\leq t \cdot \sup_{a\in q(\mathbb{R})} \lambda(q^{-1}(a) \cap A_{c,t}) \\ &\leq t\left(\frac{2D_\mu^\alpha(R_0)}{\kappa}\right)^{\frac{1}{1+r}} < \frac{\varepsilon}{2}. \end{aligned} \tag{90}$$

With (85) and (90) we finally obtain

$$\max_{a\in q(\mathbb{R})} \mu(q^{-1}(a)) \leq \max_{a\in q(\mathbb{R})} \mu(q^{-1}(a) \cap A_{c,t}) + 1 - \mu(A_{c,t}) < \varepsilon,$$

which proves (22). Now, additionally, let $i(f) > 0$. Let $a \in q(\mathbb{R})$ with $\mu(q^{-1}(a)) > 0$. By Lemma 2.3

we can assume, that $q^{-1}(a)$ is an interval. Thus we obtain

$$\lambda(q^{-1}(a) \cap I_i) = \operatorname{diam}(q^{-1}(a) \cap I_i)$$

for every $i \in \{1, \dots, m\}$. Together with $I_i \subset \operatorname{supp}(\mu)$ we deduce

$$\begin{aligned} \mu(q^{-1}(a)) \geq \mu(q^{-1}(a) \cap I_i) &= \int_{q^{-1}(a) \cap I_i} f(x) \, d\lambda(x) \\ &\geq i(f) \lambda(q^{-1}(a) \cap I_i) \\ &= i(f) \operatorname{diam}(q^{-1}(a) \cap I_i) \end{aligned} \tag{91}$$

for every $i \in \{1, \dots, m\}$. Relation (23) follows now immediately from (22) and (91). □

Proof of Lemma 5.4. Let $\alpha \in [-\infty, 0]$. Then by Lemma 2.3 for any $\gamma > 1$ there exists a quantizer q with $H_\mu^\alpha(q) \leq R$ such that each cell of q is an interval with positive $\mu-$mass (and thus $q(\mathbb{R}) = N_q$) and

$$D_\mu(q) \leq \gamma \cdot D_\mu^\alpha(R). \tag{92}$$

According to definition (2) we obtain in case of $\alpha > -\infty$ that

$$\sum_{a \in q(\mathbb{R})} \mu(q^{-1}(a))^\alpha \leq e^{(1-\alpha)R}. \tag{93}$$

We deduce

$$e^R \geq \left(\sum_{a \in q(\mathbb{R})} \mu(q^{-1}(a))(1/\mu(q^{-1}(a)))^{1-\alpha} \right)^{1/(1-\alpha)} \geq n \tag{94}$$

where n is the number of codepoints of q and the second inequality follows from Jensen's inequality applied to the concave function $x \mapsto x^{1/(1-\alpha)}$. In case of $\alpha = -\infty$ we obviously have $n < \infty$. We get

$$1 = \sum_{a \in q(\mathbb{R})} \mu(q^{-1}(a)) \geq n \cdot \min\{\mu(q^{-1}(a)) : a \in q(\mathbb{R})\} \geq n e^{-R}.$$

Hence, $n \leq e^R$ for every $\alpha \in [-\infty, 0]$. Let $\operatorname{supp}(\mu) = [c, d]$, where $-\infty < c < d < \infty$. For every $a \in q(\mathbb{R})$ let m_a denote the midpoint of $q^{-1}(a) \cap [c, d]$. As in the proof of Lemma 5.3, we obtain

$$\begin{aligned} D_\mu(q) &= \sum_{a \in q(\mathbb{R})} \int_{q^{-1}(a)} |x - a|^r f(x) \, d\lambda(x) \\ &\geq \sum_{a \in q(\mathbb{R})} i(f) \int_{q^{-1}(a) \cap [c,d]} |x - m_a|^r \, d\lambda(x) \\ &= \sum_{a \in q(\mathbb{R})} i(f) (2^r (1 + r))^{-1} \operatorname{diam}(q^{-1}(a) \cap [c, d])^{r+1}. \end{aligned}$$

Clearly (cf. (84)),

$$\frac{\mu(q^{-1}(a))}{s(f)} \leq \operatorname{diam}(q^{-1}(a) \cap [c,d]) \leq \frac{\mu(q^{-1}(a))}{i(f)}$$

for every $a \in q(\mathbb{R})$ with $\mu(q^{-1}(a)) > 0$. Thus we deduce from the convexity of $x \mapsto x^{r+1}$ that

$$\begin{aligned} D_\mu(q) &\geq i(f)(2^r(1+r))^{-1}s(f)^{-r-1} \sum_{a \in q(\mathbb{R})} (\mu(q^{-1}(a)))^{r+1} \\ &\geq i(f)(2^r(1+r))^{-1}s(f)^{-r-1} \sum_{i=1}^{n} (1/n)^{r+1} \\ &= i(f)(2^r(1+r))^{-1}s(f)^{-r-1}n^{-r}. \end{aligned}$$

Combining (92) and Proposition 5.1 (ii) we obtain

$$i(f)(2^r(1+r))^{-1}s(f)^{-r-1}n^{-r} \leq \gamma \frac{2^r}{i(f)^r} e^{-rR}. \tag{95}$$

Because $\gamma \in (1,\infty)$ was arbitrary, inequality (95) remains valid if we set $\gamma = 1$. Hence we obtain

$$\left(\frac{i(f)}{s(f)}\right)^{\frac{r+1}{r}} \left(\frac{1}{4^r(1+r)}\right)^{1/r} e^R \leq n,$$

which yields (25).

Now assume $\alpha \in (-\infty, 0)$. We will modify q such that the new quantizer is in $\mathcal{K}_R$ and it still satisfies the rate constraint, while its distortion does not exceed that of q. Let

$$p = \max\{\mu(q^{-1}(a)) : a \in q(\mathbb{R})\} > 0$$

and $a_p \in q(\mathbb{R})$ such that $\mu(q^{-1}(a_p)) = p$. If $H_\mu^\alpha(q) < R$ we can subdivide the cell $q^{-1}(a_p)$ into two cells with equal $\mu-$mass, such that the entropy increases by

$$\begin{aligned} &-\frac{1}{1-\alpha}\log\left(p^\alpha + \sum_{a \in q(\mathbb{R})\setminus\{a_p\}} \mu(q^{-1}(a))^\alpha\right) \\ &\quad + \frac{1}{1-\alpha}\log\left(2(p/2)^\alpha + \sum_{a \in q(\mathbb{R})\setminus\{a_p\}} \mu(q^{-1}(a))^\alpha\right) > 0. \end{aligned}$$

If we take the optimal quantization points for the two new cells, the new quantizer does not increase the quantization error. As long as the entropy is lower than R we repeat this procedure. Hence there exists a modified quantizer (also denoted by q) satisfying

$$e^{(1-\alpha)R} - e^{(1-\alpha)H_\mu^\alpha(q)} \leq (2^{1-\alpha} - 1)p^\alpha. \tag{96}$$

Note that (93) and (94) remain valid also for this modified quantizer. Consequently,

$$0 < Ce^R \le \mathrm{card}(q) \le e^R < \infty. \tag{97}$$

Thus we deduce

$$e^{(1-\alpha)R} \ge e^{(1-\alpha)H_\mu^\alpha(q)} = \sum_{a\in q(\mathbb{R})} \mu(q^{-1}(a))^\alpha \ge \mathrm{card}(q)\cdot p^\alpha \ge Ce^R p^\alpha,$$

which implies

$$p^\alpha \le C^{-1}e^{-R}e^{(1-\alpha)R}.$$

Together with (96) and $R > \log(\frac{2^{1-\alpha}-1}{C})$ we obtain

$$\begin{aligned} 1 &\le \frac{e^{(1-\alpha)R}}{e^{(1-\alpha)H_\mu^\alpha(q)}} = \frac{e^{(1-\alpha)R}}{e^{(1-\alpha)R} - (e^{(1-\alpha)R} - e^{(1-\alpha)H_\mu^\alpha(q)})} \\ &\le \frac{e^{(1-\alpha)R}}{e^{(1-\alpha)R} - (2^{1-\alpha}-1)p^\alpha} \\ &\le \frac{e^{(1-\alpha)R}}{e^{(1-\alpha)R} - (2^{1-\alpha}-1)C^{-1}e^{-R}e^{(1-\alpha)R}}. \end{aligned} \tag{98}$$

In view of (98) and (97) we conclude that $q \in \mathcal{K}_R$, which proves (26). □

Proof of Lemma 5.5. Recall the definition (24) of constant C. Fix $\kappa \in (0, C)$. Let $R_0 > 0$ such that $Ce^R - (m-1) \ge \kappa e^R$ for every $R \ge R_0$. According Lemma 5.3, in the definition of $D_\mu^\alpha(R)$ it suffices w.l.o.g. to consider for $R \ge R_0$ only those quantizers $q \in \mathcal{H}_R$ satisfying

$$\sup\{\mathrm{diam}(q^{-1}(a)\cap I) : a \in q(\mathbb{R})\} < \mathrm{diam}(I)/2m. \tag{99}$$

In view of Lemma 5.4 it suffices to show that for $R \ge R_0$ any quantizer $q \in \mathcal{H}_R$ that satisfies (99) can be modified such that the distortion of the new quantizer $\tilde{q}$ does not exceed that of q and it satisfies $\tilde{q} \in \mathcal{K}_R(\kappa)$ and

$$2\inf\{\mathrm{diam}(\tilde{q}^{-1}(a)\cap I) : a \in S(\tilde{q})\} \ge \inf\{\mathrm{diam}(\tilde{q}^{-1}(a)) : a \in A(\tilde{q})\}.$$

According to the upper bound (99) we always have $A(q) \ne \emptyset$. If $S(q) = \emptyset$, then the assertion is obvious. Hence, let $S(q) \ne \emptyset$. Let us assume w.l.o.g. that $\mu(q^{-1}(b)) > 0$ and that (see Lemma 2.3) $q^{-1}(b)$ is an interval for every $b \in q(\mathbb{R})$. For every $a \in S(q)$ let $\emptyset \ne N(a) \subset q(\mathbb{R}) \setminus \{a\}$ be the set of neighbor points, i.e., for every $b \in N(a)$ we have either $\sup q^{-1}(b) = \inf q^{-1}(a)$ or $\inf q^{-1}(b) = \sup q^{-1}(a)$. Due to (99) we know that $N(a) \cap S(q) = \emptyset$. Moreover, $N(a) \subset A(q)$ and $\mathrm{card}(N(a)) = 2$. Fix $i_a \in \{1, \ldots, m-1\}$ such that $q^{-1}(a) \subset I_{i_a} \cup I_{i_a+1}$. Because $a \in S(q)$, we have $\Delta_1 = \mathrm{diam}(q^{-1}(a)\cap I_{i_a}) > 0$ and $\Delta_2 = \mathrm{diam}(q^{-1}(a)\cap I_{i_a+1}) > 0$. Moreover, $\mathrm{diam}(q^{-1}(a)) = \Delta_1 + \Delta_2$.

Let $b_1 \in N(a)$ such that $\inf(q^{-1}(a)) = \sup(q^{-1}(b_1))$ and let $b_2 \in N(a)$ such that $\inf(q^{-1}(b_2)) = \sup(q^{-1}(a))$. Next we will show that

$$\inf(q^{-1}(a)) + \frac{\Delta_1}{2} \leq a \leq \inf(q^{-1}(a)) + \Delta_1 + \frac{\Delta_2}{2}. \tag{100}$$

To see this, one recognizes that a has to be optimal for $\mu(\cdot|q^{-1}(a))$. As a consequence (see, e.g., [9, Lemma 2.6 (a)]), $a \in [\inf(q^{-1}(a)), \sup(q^{-1}(a))]$. Moreover, a has to be a stationary point (see [9, Lemma 2.5]), which yields

$$\int_{[\inf(q^{-1}(a)),a]} |x-a|^{r-1}\, d\mu(x) = \int_{[a,\sup(q^{-1}(a))]} |x-a|^{r-1}\, d\mu(x). \tag{101}$$

Now let us assume that the first inequality in (100) does not hold. Hence,

$$a < \inf(q^{-1}(a)) + \Delta_1/2. \tag{102}$$

Note that $\sup(I_{i_a}) = \inf q^{-1}(a) + \Delta_1$ and that $\sup(I_{i_a}) + \Delta_2 = \sup(q^{-1}(a))$. From (101) and $\Delta_2 > 0$ we get

$$\int_{[\inf(q^{-1}(a)),a]} |x-a|^{r-1}\, d\mu(x) \quad > \quad \int_{[a,\inf q^{-1}(a)+\Delta_1]} |x-a|^{r-1}\, d\mu(x). \tag{103}$$

Because the density of μ is constant on $[\inf(q^{-1}(a)), \inf(q^{-1}(a)) + \Delta_1]$ we obtain from (103) that $a > \inf(q^{-1}(a)) + \Delta_1/2$, which contradicts (102). Thus we have proved the left inequality in (100). Similarly, we deduce from $\Delta_1 > 0$ and (101) the right inequality in (100).

Recall that μ has constant density on $q^{-1}(b_i)$; $i = 1, 2$. Again by stationarity (101) we obtain

$$b_1 = \inf q^{-1}(a) - \mathrm{diam}(q^{-1}(b_1))/2 \tag{104}$$

and

$$b_2 = \inf q^{-1}(a) + \Delta_1 + \Delta_2 + \mathrm{diam}(q^{-1}(b_2))/2.$$

Let $\Delta = \Delta_1 + \Delta_2$. Next we show that w.l.o.g. we can assume $2\Delta \geq \min(\mathrm{diam}(q^{-1}(b_1)), \mathrm{diam}(q^{-1}(b_2)))$. Assume to the contrary that

$$2\Delta < \min(\mathrm{diam}(q^{-1}(b_1)), \mathrm{diam}(q^{-1}(b_2))). \tag{105}$$

Then we have $\mathrm{diam}(q^{-1}(b_1)) > 2\Delta > 2\Delta_1 + \Delta_2$, and applying (104) we get

$$\inf q^{-1}(a) - b_1 > \inf q^{-1}(a) + \Delta_1 + \frac{\Delta_2}{2} - \inf q^{-1}(a)$$

Hence, (100) implies

$$\inf q^{-1}(a) - b_1 > a - \inf q^{-1}(a). \tag{106}$$

Similarly we obtain

$$b_2 - \sup q^{-1}(a) > \sup q^{-1}(a) - a. \tag{107}$$

In view of (105) and by the definition of μ we have

$$\frac{\operatorname{diam}(I)}{m}\,\mu(q^{-1}(b_1)) = \operatorname{diam}(q^{-1}(b_1)) \cdot s_{i_a} > 2\Delta \cdot s_{i_a}$$

and

$$\frac{\operatorname{diam}(I)}{m}\,\mu(q^{-1}(b_2)) = \operatorname{diam}(q^{-1}(b_2)) \cdot s_{i_a+1} > 2\Delta \cdot s_{i_a+1}.$$

Moreover,

$$\frac{\operatorname{diam}(I)}{m}\,\mu(q^{-1}(a)) = \Delta_1 s_{i_a} + \Delta_2 s_{i_a+1} < 2\Delta \max\{s_{i_a}, s_{i_a+1}\}.$$

Thus we obtain

$$\mu(q^{-1}(a)) < \max(\mu(q^{-1}(b_1)), \mu(q^{-1}(b_2))) \tag{108}$$

as long as (105) holds. Thus, in view of (106) and (107), we can modify q by increasing the codecell $q^{-1}(a)$, which yields a reduction of the quantization error and a non-increasing entropy of q (due to $\alpha < 0$, as long as (108) holds, the entropy is a non-decreasing function of the left endpoint of the cell $q^{-1}(a)$ and a non-increasing function of the right endpoint of $q^{-1}(a)$). The codecell can be expanded this way until $2\Delta = \min(\operatorname{diam}(q^{-1}(b_1)), \operatorname{diam}(q^{-1}(b_2)))$ holds. Note that independent of this modification q remains an element of $\mathcal{H}_R$. Thus we can assume w.l.o.g. that

$$2\Delta \geq \min(\operatorname{diam}(q^{-1}(b_1)), \operatorname{diam}(q^{-1}(b_2))). \tag{109}$$

If $q \in \mathcal{K}_R(\kappa)$, then the proof is finished. Hence, let us assume that $q \notin \mathcal{K}_R(\kappa)$. We will show that q can always be modified such that the new quantizer belongs to $\mathcal{K}_R(\kappa)$ and still satisfies relation (109). We proceed as in the proof of relation (26). Let

$$W(q) = \{a \in q(\mathbb{R}) : \mu(q^{-1}(a)) = \max\{\mu(q^{-1}(b)) : b \in A(q)\}\}.$$

We subdivide one by one the cells $q^{-1}(a)$ with $a \in W(q)$ and $p = \mu(q^{-1}(a))$ as in the proof of (26) in Lemma 5.4. Note, that the entropy of the quantizer will exceed any given bound if we repeat the subdivision process enough times. We stop this process with a quantizer $\tilde{q}$ that satisfies relation (96). Now recall that $Ce^R - (m-1) \geq \kappa e^R$ if $R \geq R_0$ by the definition at the beginning of the proof. Thus, with $p = \mu(\tilde{q}^{-1}(a))$, we have

$$\begin{aligned} e^{(1-\alpha)R} &\geq e^{(1-\alpha)H_\mu^\alpha(\tilde{q})} \geq (\operatorname{card}(\tilde{q}(\mathbb{R})) - (m-1))p^\alpha \\ &\geq (Ce^R - (m-1))p^\alpha \geq \kappa e^R p^\alpha. \end{aligned}$$

Now the inequality $e^{(1-\alpha)R} \geq me^R p^\alpha$ allows us to perform steps identical to the ones in the chain of inequalities (98) and we obtain that the quantizer belongs to $\mathcal{K}_R(\kappa)$. Obviously, (109) is still in force for $\tilde{q}$ and the proof is complete. $\square$

Appendix C

Proof of Proposition 6.2. For every $i \in \{1, \ldots, m\}$ choose a quantizer $q_i \in \mathcal{Q}$ for μ_i with $H^\alpha_{\mu_i}(q_i) \leq R_i$. Let

$$J_i = \{a \in q_i(\mathbb{R}) : \mu(q_i^{-1}(a) \cap A_i) > 0\}.$$

Let $I_i \subset \mathbb{N}$ be an index set of the same cardinality as J_i and for every $k \in I_i$ choose $a_{i,k} \in q_i(\mathbb{R})$ such that $J_i = \{a_{i,k} : k \in I_i\}$. Let

$$N = \mathbb{R} \setminus \cup_{i=1}^m \cup_{k \in I_i} q_i^{-1}(a_{i,k}) \cap A_i.$$

Note that $\mu(N) = 0$. Now we define the quantizer q by the codecells

$$\{N\} \cup \{q_i^{-1}(a_{i,k}) \cap A_i : i = 1, \ldots, m; k \in I_i\}$$

and corresponding codepoints

$$\{0\} \cup \{a_{i,k} : i = 1, \ldots, m; k \in I_i\}.$$

Note that despite our general assumption, the codepoints now are not necessarily distinct. Recall the convention $0^0 = 0$. Since $\mu(N) = 0$, the definition of $H^\alpha_\mu(q)$ yields

$$\begin{aligned} H^\alpha_\mu(q) &= \frac{1}{1-\alpha} \log \left(\sum_{i=1}^m \sum_{k \in I_i} \mu(q_i^{-1}(a_{i,k}) \cap A_i)^\alpha \right) \\ &= \frac{1}{1-\alpha} \log \left(\sum_{i=1}^m s_i^\alpha \sum_{a \in q_i(\mathbb{R})} \mu_i(q_i^{-1}(a))^\alpha \right) \\ &= \frac{1}{1-\alpha} \log \left(\sum_{i=1}^m s_i^\alpha e^{(1-\alpha)H^\alpha_{\mu_i}(q_i)} \right). \end{aligned}$$

Since $H^\alpha_{\mu_i}(q_i) \leq R_i$, we obtain in both cases ($\alpha < 1$ and $\alpha > 1$) that

$$H^\alpha_\mu(q) \leq \frac{1}{1-\alpha} \log \left(\sum_{i=1}^m s_i^\alpha e^{(1-\alpha)R_i} \right).$$

Now it is easy to check that $H_\mu^\alpha(q) \le R$ is satisfied if either (28) or (29) holds. Further we deduce

$$\begin{aligned} D_\mu^\alpha(R) &\le D_\mu(q) = \int |x - q(x)|^r \, d\mu(x) \\ &= \sum_{i=1}^m s_i \int_{A_i} |x - q_i(x)|^r \, d\mu_i(x) = \sum_{i=1}^m s_i D_{\mu_i}(q_i). \end{aligned}$$

Taking the infimum on the right hand side of above inequality yields the assertion. □

Proof of Lemma 6.3. From Definition 6.1 we have $s \in (0,1)$. Let $R \ge 0$ and $\delta > 0$. Let $q \in \mathcal{Q}$ with $H_\mu^\alpha(q) \le R$ and $\delta + D_\mu^\alpha(R) \ge D_\mu(q)$. We obtain

$$\delta + D_\mu^\alpha(R) \ge D_\mu(q) \ge s \int |x - q(x)|^r \, d\mu_{i_0}(x). \tag{110}$$

Since $\alpha \in [0,1)$, we deduce

$$\begin{aligned} R &\ge H_\mu^\alpha(q) = \frac{1}{1-\alpha} \log\left(\sum_{a \in q(\mathbb{R})} \left(\sum_{i=1}^m s_i \mu_i(q^{-1}(a)) \right)^\alpha \right) \\ &\ge \frac{1}{1-\alpha} \log\Bigg(\sum_{a \in q(\mathbb{R})} \left(s\mu_{i_0}(q^{-1}(a)) \right)^\alpha \Bigg) \\ &= \frac{\alpha}{1-\alpha} \log(s) + H_{\mu_{i_0}}^\alpha(q). \end{aligned}$$

Because δ was arbitrary we get from (110) that

$$D_\mu^\alpha(R) \ge s D_{\mu_{i_0}}^\alpha \left(R - \frac{\alpha}{1-\alpha} \log(s) \right),$$

which yields

$$\begin{aligned} e^{rR} D_\mu^\alpha(R) &\ge s e^{r\left(\frac{\alpha}{1-\alpha}\log(s)\right)} e^{r\left(R - \frac{\alpha}{1-\alpha}\log(s)\right)} D_{\mu_{i_0}}^\alpha \left(R - \frac{\alpha}{1-\alpha} \log(s) \right) \\ &= s^{a_1 a_2} e^{r\left(R - \frac{\alpha}{1-\alpha}\log(s)\right)} D_{\mu_{i_0}}^\alpha \left(R - \frac{\alpha}{1-\alpha} \log(s) \right) \end{aligned}$$

and therefore proves (31).

Now let $\alpha \in [0, r+1) \setminus \{1\}$ and fix $R_0 > 0$, such that

$$R_0 \ge \max\{-\log(t_i) : i = 1, \dots, m\}.$$

For any $R > R_0$ let $R_i = R + \log(t_i) > 0$, $i = 1, \ldots, m$. We obtain

$$\sum_{i=1}^{m} s_i^{\alpha} e^{(1-\alpha)R_i} = e^{(1-\alpha)R}, \tag{111}$$

if $\alpha \in [0, r+1) \setminus \{1\}$. Indeed, (111) is equivalent to $\sum_{i=1}^{m} s_i^{\alpha} t_i^{1-\alpha} = 1$. But this equation is satisfied by the definition of t_i. Applying Proposition 6.2 we obtain

$$D_{\mu}^{\alpha}(R) \leq s D_{\mu_{i_0}}^{\alpha}(R_{i_0}) + \sum_{i=1; i \neq i_0}^{m} s_i D_{\mu_i}^{\alpha}(R_i).$$

Thus we can compute

$$\begin{aligned} e^{rR} D_{\mu}^{\alpha}(R) &\leq e^{rR} s D_{\mu_{i_0}}^{\alpha}(R_{i_0}) + \sum_{i=1; i \neq i_0}^{m} e^{rR} s_i D_{\mu_i}^{\alpha}(R_i) \\ &= e^{r(R-R_{i_0})} s e^{rR_{i_0}} D_{\mu_{i_0}}^{\alpha}(R_{i_0}) + \sum_{i=1; i \neq i_0}^{m} e^{r(R-R_i)} s_i e^{rR_i} D_{\mu_i}^{\alpha}(R_i) \\ &= s t_{i_0}^{-r} e^{rR_{i_0}} D_{\mu_{i_0}}^{\alpha}(R_{i_0}) + \sum_{i=1; i \neq i_0}^{m} s_i t_i^{-r} e^{rR_i} D_{\mu_i}^{\alpha}(R_i). \end{aligned} \tag{112}$$

Because all terms in (112) are nonnegative we obtain (32). □

Appendix D

Lemma D.1. *Let $m \in \mathbb{N}$ and $\sum_{i=1}^{m} s_i = 1$ with $s_i > 0$ for every $i \in \{1, \ldots, m\}$. Let the probability measure μ be supported on a bounded interval I such that $\mu = \sum_{i=1}^{m} s_i U(I_i)$ where the I_i are intervals of equal length $\lambda(I)/m$ that partition I. Let $\alpha \in (-\infty, 0)$ and $(R_n)_{n \in \mathbb{N}}$ be an increasing sequence of positive numbers such that $R_n \to \infty$ as $n \to \infty$. Then for every sequence $(q_n)_{n \in \mathbb{N}}$ of quantizers with $q_n \in \mathcal{G}_{R_n}$, relation (48) holds.*

Proof. Recall from (27) the definition of $A(q)$ and $S(q)$. For any $n \in \mathbb{N}$

$$\begin{aligned} 1 &\leq \frac{e^{(1-\alpha)H_{\mu}^{\alpha}(q_n)}}{\sum_{a \in A(q_n)} \mu(q_n^{-1}(a))^{\alpha}} \\ &= \frac{\sum_{a \in A(q_n)} \mu(q_n^{-1}(a))^{\alpha} + \sum_{a \in S(q_n)} \mu(q_n^{-1}(a))^{\alpha}}{\sum_{a \in A(q_n)} \mu(q_n^{-1}(a))^{\alpha}} \\ &\leq 1 + \frac{\operatorname{card}(S(q_n)) \cdot \sup\{\mu(q_n^{-1}(a))^{\alpha} : a \in S(q_n)\}}{\sum_{a \in A(q_n)} \mu(q_n^{-1}(a))^{\alpha}} \\ &\leq 1 + (m-1) \frac{\sup\{\mu(q_n^{-1}(a))^{\alpha} : a \in S(q_n)\}}{\sum_{a \in A(q_n)} \mu(q_n^{-1}(a))^{\alpha}} \end{aligned}$$

$$= \quad 1+(m-1)\frac{(\inf\{\mu(q_n^{-1}(a)) : a \in S(q_n)\})^\alpha}{\sum_{a\in A(q_n)}\mu(q_n^{-1}(a))^\alpha}. \tag{113}$$

Now let

$$h_1 = \min\left\{\frac{s_i}{\lambda(I)/m} : i \in \{1,\ldots,m\}\right\} > 0$$

and

$$h_2 = \max\left\{\frac{s_i}{\lambda(I)/m} : i \in \{1,\ldots,m\}\right\} > 0.$$

Since $q_n \in \mathcal{G}_{R_n}$, we have

$$\begin{aligned} 1 &\leq 1+(m-1)(h_1/2)^\alpha\frac{(\min\{\mathrm{diam}(q_n^{-1}(a)) : a \in A(q_n)\})^\alpha}{\sum_{a\in A(q_n)}\mu(q_n^{-1}(a))^\alpha} \\ &\leq 1+(m-1)(h_1/2h_2)^\alpha\frac{(\min\{\mathrm{diam}(q_n^{-1}(a)) : a \in A(q_n)\})^\alpha}{\sum_{a\in A(q_n)}\mathrm{diam}(q_n^{-1}(a))^\alpha}. \end{aligned} \tag{114}$$

Fix $i = i(n) \in \{1,\ldots,m\}$ and $b \in A(q_n) \cap I_i$ such that

$$\mathrm{diam}(q_n^{-1}(b)) = \min\{\mathrm{diam}(q_n^{-1}(a)) : a \in A(q_n)\}. \tag{115}$$

From Proposition 7.2 and by [9, Example 5.5] we know that all codecells $q_n^{-1}(a)$ with $a \in A(q_n) \cap I_i$ can be assumed to have equal length. Because $q_n \in \mathcal{G}_{R_n} \subset \mathcal{K}_{R_n}$ we obtain $\lim_{n\to\infty} H_\mu^\alpha(q_n) = \infty$. In view of (115) we thus get $\mathrm{card}(A(q_n) \cap I_i) \to \infty$ as $n \to \infty$. From (113) and (114) we deduce

$$\begin{aligned} 1 &\leq \frac{e^{(1-\alpha)H_\mu^\alpha(q_n)}}{\sum_{a\in A(q_n)}\mu(q_n^{-1}(a))^\alpha} \leq 1+\frac{(m-1)(h_1/2h_2)^\alpha(\mathrm{diam}(q_n^{-1}(b)))^\alpha}{\sum_{a\in A(q_n)\cap I_i}\mathrm{diam}(q_n^{-1}(a))^\alpha} \\ &= 1+\frac{(m-1)(h_1/2h_2)^\alpha}{\mathrm{card}(A(q_n)\cap I_i)} \to 1 \text{ as } n \to \infty. \end{aligned} \tag{116}$$

Again from $q_n \in \mathcal{G}_{R_n} \subset \mathcal{K}_{R_n}$ we have $\lim_{n\to\infty} e^{(1-\alpha)(R_n - H_\mu^\alpha(q_n))} = 1$, which yields together with (116) the assertion. $\square$

Let $m \geq 2$ and $s_1,\ldots,s_m \in (0,1)^m$ with $\sum_{i=1}^m s_i = 1$. For $(v_1,\ldots,v_m) \in (0,\infty)^m$ and $\alpha \in (-\infty,\infty)\setminus\{1\}$ we define

$$F(v_1,\ldots,v_m) = \sum_{i=1}^m s_i v_i^{-r}$$

and set $t_i = s_i^{1/a_2}\left(\sum_{j=1}^m s_j^{a_1}\right)^{-\frac{1}{1-\alpha}}$, $i = 1,\ldots,m$ as in (30).

Lemma D.2. *If* $\alpha \in (-\infty,1)$*, then*

$$F(t_1,\ldots,t_m) = \inf\{F(v_1,\ldots,v_m) : (v_1,\ldots,v_m) \in (0,\infty)^m;\ \sum_{i=1}^m s_i^\alpha v_i^{1-\alpha} = 1\}.$$

Proof. Let $x_i = s_i^{\alpha} v_i^{1-\alpha}$. We calculate

$$v_i = (x_i s_i^{-\alpha})^{\frac{1}{1-\alpha}}$$

and

$$\begin{aligned} F(v_1, \ldots, v_m) &= \sum_{i=1}^{m} s_i (x_i s_i^{-\alpha})^{\frac{-r}{1-\alpha}} \\ &= \sum_{i=1}^{m} s_i^{\frac{1-\alpha+\alpha r}{1-\alpha}} x_i^{-\frac{r}{1-\alpha}} =: G(x_1, \ldots, x_m). \end{aligned}$$

Applying [9, Lemma 6.8] we deduce that G attains its minimum on $(0, \infty)^m$ subject to the constraint $\sum_{i=1}^{m} x_i = 1$ at the point $(y_1, \ldots, y_m)$ with

$$y_i = \frac{\left(s_i^{\frac{1-\alpha+\alpha r}{1-\alpha}}\right)^{\frac{1}{1+\frac{r}{1-\alpha}}}}{\sum_{j=1}^{m} \left(s_j^{\frac{1-\alpha+\alpha r}{1-\alpha}}\right)^{\frac{1}{1+\frac{r}{1-\alpha}}}} = \frac{s_i^{a_1}}{\sum_{j=1}^{m} s_j^{a_1}}$$

for every $i \in \{1, \ldots, m\}$. Hence, F attains its minimum subject to the constraint $\sum_{i=1}^{m} s_i^{\alpha} v_i^{1-\alpha} = 1$ at the point $(w_1, \ldots, w_m)$ with $w_i = (y_i s_i^{-\alpha})^{\frac{1}{1-\alpha}}$ for every $i \in \{1, \ldots, m\}$. We deduce

$$w_i^{1-\alpha} = \frac{s_i^{\frac{1-\alpha+\alpha r}{1-\alpha+r}} s_i^{-\alpha}}{\sum_{j=1}^{m} \left(s_j^{\frac{1-\alpha+\alpha r}{1-\alpha}}\right)^{\frac{1}{1+\frac{r}{1-\alpha}}}} = \frac{s_i^{\frac{(1-\alpha)^2}{1-\alpha+r}}}{\sum_{j=1}^{m} s_j^{a_1}}$$

which yields $w_i = t_i$. □

Acknowledgments

The authors would like to thank two anonymous reviewers for their detailed and constructive comments.

References

[1] J. Aczél and Z. Daróczy, *On Measures of Information and Their Characterizations*. London: Academic Press, 1975.

[2] M. B. Baer, *Coding for General Penalties*. Ph.D. Thesis, Stanford University, June 2003.

[3] W. R. Bennett, "Spectra of quantized signals," *Bell. Syst. Tech. J.*, vol. 27, pp. 446–472, 1948.

[4] P. Billingsley, *Probability and Measure*. New York: Wiley, 2nd ed., 1986.

[5] J. A. Bucklew and G. L. Wise, "Multidimensional asymptotic quantization theory with rth power distortion measures," *IEEE Trans. Inform. Theory*, vol. 28, pp. 239–247, Mar. 1982.

[6] L. L. Campbell, “A coding theorem and Rényi’s entropy,” *Information and Control*, vol. 8, pp. 423–429, 1965.

[7] D. L. Cohn, *Measure Theory*. Boston: Birkhäuser, 1980.

[8] I. Csiszár, “Generalized entropy and quantization problems,” in *Transactions of the Sixth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes,* (Prague), pp. 29–35, Akademia, 1973.

[9] S. Graf and H. Luschgy, *Foundations of Quantization for Probability Distributions*. Berlin, Heidelberg: Springer Verlag, 2000.

[10] R. M. Gray, T. Linder, and J. T. Gill, “Lagrangian vector quantization with combined entropy and codebook size constraints,” *IEEE Trans. Inform. Theory*, vol. 54, pp. 2220–2242, May 2008.

[11] R. M. Gray, T. Linder, and J. Li, “A Lagrangian formulation of Zador’s entropy-constrained quantization theorem,” *IEEE Trans. Inform. Theory*, vol. 48, pp. 695–707, Mar. 2002.

[12] R. M. Gray and D. L. Neuhoff, “Quantization,” *IEEE Trans. Inform. Theory,* (Special Commemorative Issue), vol. 44, pp. 2325–2383, Oct. 1998.

[13] A. György and T. Linder, “Optimal entropy-constrained scalar quantization of a uniform source,” *IEEE Trans. Inform. Theory*, vol. 46, pp. 2704–2711, Nov. 2000.

[14] P. Harremoës, “Joint range of Rényi entropies,” *Kybernetika*, vol. 45, pp. 901–911, 2009.

[15] A. S. Holevo, “The additivity problem in quantum information theory,” *Proc. of the International Congress of Mathematicians*, (Madrid, Spain), pp. 999–1018, 2006.

[16] F. Jelinek, “Buffer overflow in variable length coding of fixed rate sources,” *IEEE Trans. Inform. Theory*, vol. 14, pp. 490–501, May 1968.

[17] R. Jenssen, “Information theoretic learning and kernel methods,” in *Information Theory and Statistical Learning*, F. Emmert-Streib and M. Dehmer, eds., pp. 209–230, Springer, New York, 2009.

[18] W. Kreitmeier, “Error bounds for high-resolution quantization with Rényi-α-entropy constraints,” *Acta Math. Hungar.*, vol. 127, pp. 34–51, 2010.

[19] W. Kreitmeier, “Optimal quantization for the one-dimensional uniform distribution with Rényi-α-entropy constraints,” *Kybernetika*, vol. 46, pp. 96–113, 2010.

[20] W. Kreitmeier, “Optimal vector quantization in terms of Wasserstein distance,” *Journal of Multivariate Analysis,* doi:10.1016/j.jmva.2011.04.005, vol. 102, no. 8, pp. 1225–1239, Sep. 2011.

[21] A. Krishnamacharia, V. moy Mandalb, Karmeshu, ”Study of DNA binding sites using the Rényi parametric entropy measure,” *Journal of Theoretical Biology*, vol. 227, no. 3, pp. 429–436, 2004.

[22] T. Linder, “On asymptotically optimal companding quantization,” *Probl. Control Inf. Theory*, vol. 20, no. 6, pp. 383–393, 1991.

[23] H. Luschgy and G. Pagès, “Functional quantization rate and mean regularity of processes with an application to Lévy processes,” *Ann. Appl. Probab.*, vol. 18, pp. 427–469, 2008.

[24] P. Nath, “On a coding theorem connected with Rényi’s entropy,” *Information and Control*, vol. 29, pp. 234–242, 1975.

[25] J. N. Pierce, “Asymptotic quantizing error for unbounded random variables,” *IEEE Trans. Inform. Theory*, vol. 16, pp. 81–83, Jan. 1970.

[26] A. Rényi, “On measures of entropy and information.” in *Proc. 4th Berkeley Symp. Math. Stat. Probab.*, vol. 1, pp. 547–561, 1960.

[27] A. Rényi, “Some fundamental problems of information theory (in Hungarian),” *Magyar Tud. Akad. Mat. Fiz. Oszt. Közl*, vol. 10, pp. 251–282, 1960.

[28] H. L. Royden, *Real Analysis*. New York: Macmillan, 1968.

[29] A. Sagna, *Méthodes de quantification optimale avec applications à la Finance,*. PhD thesis, Université Paris VI, 2008.

[30] J. Serrin and D. E. Varberg, “A general chain rule for derivatives and the change of variables formula for the Lebesgue integral,” *Am. Math. Mon.*, vol. 76, pp. 514–520, May 1969.

[31] P. Zador, *Development and evaluation of procedures for quantizing multivariate distributions*. PhD thesis, Stanford Univ., 1963. Univ. Microfilm no. 64-9855.